

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein -- protein search, using sw model

Run on: June 5, 2003, 16:12:50 ; Search time 27 seconds
(without alignments)
545.959 Million cell updates/sec

Title: US-09-869-414a-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLMGAGVLPAAH.....CLRLRQHQHDFADDISLKK 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

1 number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2656	99.7	501	4	US-09-548-372D-4
2	2656	99.7	501	4	US-09-548-367D-4
3	2650	99.5	501	4	US-09-009-191-2
4	2582	96.9	501	4	US-09-713-158-2
5	2567	96.4	501	4	US-09-548-372D-8
6	2567	96.4	501	4	US-09-548-372D-8
7	2506.5	94.1	476	4	US-09-548-372D-6
8	2506.5	94.1	476	4	US-09-548-367D-6
9	2420.5	90.9	476	4	US-09-548-372D-73
10	2420.5	90.9	476	4	US-09-548-367D-73
11	2397	90.0	453	4	US-09-548-372D-30
12	2397	90.0	453	4	US-09-548-367D-30
13	2397	90.0	459	4	US-09-548-372D-32
14	2397	90.0	459	4	US-09-548-367D-32
15	2315	86.9	774	4	US-09-009-191-4
16	2291.5	86.0	446	4	US-09-548-372D-22
17	2291.5	86.0	446	4	US-09-548-367D-22
18	2288	85.9	433	4	US-09-548-372D-26
19	2288	85.9	433	4	US-09-548-367D-26
20	2288	85.9	459	4	US-09-548-372D-24
21	2288	85.9	459	4	US-09-548-367D-24
22	2247.5	84.4	428	4	US-09-548-372D-51
23	2247.5	84.4	428	4	US-09-548-367D-51
24	2247.5	84.4	434	4	US-09-548-372D-53
25	2247.5	84.4	434	4	US-09-548-367D-53
26	2104	79.0	425	4	US-09-548-372D-28
27	2104	79.0	425	4	US-09-548-367D-28

28	1173.5	44.1	518	3	US-08-999-723-2	Sequence 2, Appli
29	1173.5	44.1	518	4	US-09-434-427-2	Sequence 2, Appli
30	1173.5	44.1	518	4	US-09-548-372D-2	Sequence 2, Appli
31	1173.5	44.1	518	4	US-09-548-367D-2	Sequence 2, Appli
32	1150	43.2	514	4	US-09-717-432-2	Sequence 2, Appli
33	1150	43.2	514	4	US-09-912-484-2	Sequence 2, Appli
34	301.5	11.3	396	1	US-08-208-007A-13	Sequence 13, Appl
35	301.5	11.3	396	4	US-09-032-523-9	Sequence 9, Appli
36	301.5	11.3	396	4	US-08-915-095A-13	Sequence 13, Appl
37	301.5	11.3	396	4	US-08-798-096-13	Sequence 13, Appl
38	301.5	11.3	396	4	US-08-798-095A-13	Sequence 13, Appl
39	300.5	11.3	412	1	US-08-208-007A-12	Sequence 12, Appl
40	300.5	11.3	412	4	US-08-974-691-4	Sequence 4, Appli
41	300.5	11.3	412	4	US-08-915-095A-12	Sequence 12, Appl
42	300.5	11.3	412	4	US-08-798-096-12	Sequence 12, Appl
43	300.5	11.3	412	4	US-08-798-095A-12	Sequence 12, Appl
44	282.5	10.6	458	6	5217891-15	Patent No. 5217891
45	276.5	10.4	427	2	US-08-846-021A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-548-372D-4
; Sequence 4, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-4

Query Match	99.7%	Score 2656;	DB 4;	Length 501;
Best Local Similarity	99.8%	Pred. No. 2.9e-267;		
Mismatches	0;	Mismatches	1;	Indels
Gaps	0;	Gaps	0;	
Qy	1	MAQALPWLMLMGAGVLPAAHGTQHGI	RLPLRSLGGLGAPLGLRLPRETDEEPEEGRGGSF	60
Db	1	MAQALPWLMLMGAGVLPAAHGTQHGI	RLPLRSLGGLGAPLGLRLPRETDEEPEEGRGGSF	60
Qy	61	VEWVDNLGRSGGQYVYVEMTVGSPPTL	NLIVDTGSSNFVAGAAPHPFLHRYQRLSST	120
Db	61	VEWVDNLGRSGGQYVYVEMTVGSPPTL	NLIVDTGSSNFVAGAAPHPFLHRYQRLSST	120
Qy	121	YRDLRKGVVYPYTGCKWEGELGTDLVSI	PHGPNVTVRANTAAITESDKFFINGSNWEGIL	180
Db	121	YRDLRKGVVYPYTGCKWEGELGTDLVSI	PHGPNVTVRANTAAITESDKFFINGSNWEGIL	180
Qy	181	GLAYAEIARPDSDLPEPFDSLVKQTHVP	NFLFSLHLCGAGPPLNQSEVLASVGGSMIIIGI	240
Db	181	GLAYAEIARPDSDLPEPFDSLVKQTHVP	NFLFSLHLCGAGPPLNQSEVLASVGGSMIIIGI	240
Qy	241	DHSLYTGSLWYTPIRREMYVEVIVRVEIN	QODLKMCKEYNDKSIDVSGTTNLRLPKK	300
Db	241	DHSLYTGSLWYTPIRREMYVEVIVRVEIN	QODLKMCKEYNDKSIDVSGTTNLRLPKK	300

361 ILPOQYLRPVEDVATSDCCYKFAISQSSTGTVMGAVIMEGFYVVFDRARRKRTGFAVSAC 420
361 ILPOQYLRPVEDVATSDCCYKFAISQSSTGTVMGAVIMEGFYVVFDRARRKRTGFAVSAC 420
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAAICALFMLPCLCLMVCOW 480
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAAICALFMLPCLCLMVCOW 480
481 RCLRCRQOHHDDFADDDISLLK 501
481 RCLRCRQOHHDDFADDDISLLK 501

RESULT 3
US-09-009-191-2
; Sequence 2, Application US/09009191
; Patent No. 6319689
; GENERAL INFORMATION:
; APPLICANT: POWELL, DAVID
; APPLICANT: CHAPMAN, CONRAD
; APPLICANT: MURPHY, KAY
; APPLICANT: SMITH, TRUDI
; TITLE OF INVENTION: ASP2
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,191
; FILING DATE: 20-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 9701684.4
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70368
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-009-191-2

Query Match 99.5%; Score 2650; DB 4; Length 501;
Best Local Similarity 99.6%; Pred. No. 1.2e-266;
Matches 499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAQALPWLMLLMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSF 60
DB 1 MAQALPWLMLLMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSF 60
QY 61 VEMVDNLRKSGQGYVYEMTVGSPPTNLNLDVTGSSNFVAGAAPHPFLHRYQRLSST 120
DB 61 VEMVDNLRKSGQGYVYEMTVGSPPTNLNLDVTGSSNFVAGAAPHPFLHRYQRLSST 120
QY 121 YRDLRGVYVPYTGKWEGLTDLVSIHPGPNVTVRANIAAITESDKFFINGSNWEGIL 180

301 VFEAAVKSIIKAASSTKFPDGFVLGEQVLCWQAGTTPWNIFFVISLYLMGEVTVNSFRIT 360
301 VFEAAVKSIIKAASSTKFPDGFVLGEQVLCWQAGTTPWNIFFVISLYLMGEVTVNSFRIT 360
361 ILPOQYLRPVEDVATSDCCYKFAISQSSTGTVMGAVIMEGFYVVFDRARRKRTGFAVSAC 420
361 ILPOQYLRPVEDVATSDCCYKFAISQSSTGTVMGAVIMEGFYVVFDRARRKRTGFAVSAC 420
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAAICALFMLPCLCLMVCOW 480
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAAICALFMLPCLCLMVCOW 480
481 RCLRCRQOHHDDFADDDISLLK 501
481 RCLRCRQOHHDDFADDDISLLK 501

RESULT 2
US-548-367D-4
; Sequence 4, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-548-367D-4

Query Match 99.7%; Score 2656; DB 4; Length 501;
Best Local Similarity 99.8%; Pred. No. 2.9e-267;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 MAQALPWLMLLMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSF 60
DB 1 MAQALPWLMLLMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSF 60
QY 61 VEMVDNLRKSGQGYVYEMTVGSPPTNLNLDVTGSSNFVAGAAPHPFLHRYQRLSST 120
DB 61 VEMVDNLRKSGQGYVYEMTVGSPPTNLNLDVTGSSNFVAGAAPHPFLHRYQRLSST 120
QY 121 YRDLRGVYVPYTGKWEGLTDLVSIHPGPNVTVRANIAAITESDKFFINGSNWEGIL 180
DB 121 YRDLRGVYVPYTGKWEGLTDLVSIHPGPNVTVRANIAAITESDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLEPFDSLVKQTHVPNLFSLHLCGAGFPLNQLSEVLASVGGSMIIGI 240
DB 181 GLAYAEIARPDSDLEPFDSLVKQTHVPNLFSLHLCGAGFPLNQLSEVLASVGGSMIIGI 240
QY 241 DHSLYTGSLSWTPTIRREWYEVIIIVRVEINGODLKMCKEYNDKSIIVDSGTTNLRPKK 300
DB 241 DHSLYTGSLSWTPTIRREWYEVIIIVRVEINGODLKMCKEYNDKSIIVDSGTTNLRPKK 300
QY 301 VFEAAVKSIIKAASSTKFPDGFVLGEQVLCWQAGTTPWNIFFVISLYLMGEVTVNSFRIT 360
DB 301 VFEAAVKSIIKAASSTKFPDGFVLGEQVLCWQAGTTPWNIFFVISLYLMGEVTVNSFRIT 360

Db 121 YRDLKRGVYPTQGWEGELGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLPPFDLSLVKQTHVNLPSLHLCGAGFPLNOSSEVLASVGGSMIIGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVNLPSLHLCGAGFPLNOSSEVLASVGGSMIIGI 240
Qy 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYVDSGTTNLRPKK 300
Db 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYVDSGTTNLRPKK 300
Qy 301 VFEAAVKSIIKAASSTKFPDGFNLGEQLVWCQAGTTPWNIIPVLSLYLMGEVTVNOSFRIT 360
Db 301 VFEAAVKSIIKAASSTKFPDGFNLGEQLVWCQAGTTPWNIIPVLSLYLMGEVTVNOSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
Qy 421 HVHDEFRTAAVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQOW 480
Db 421 HVHDEFRTAAVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQOW 480
Qy 481 RCLRCRQOHHDFADDISLLK 501
Db 481 RCLRCRQOHHDFADDISLLK 501

RESULT 4

US-09-713-158-2

; Sequence 2, Application US/09713158

; Patent No. 6361975

; GENERAL INFORMATION:

; APPLICANT: ZHU, YUAN

; APPLICANT: LI, XIAOTONG

; APPLICANT: POWELL, DAVID J.

; APPLICANT: CHRISTIE, GARY

; TITLE OF INVENTION: MOUSE ASPARTIC SECRETASE-2 (MASP-2)

; FILE REFERENCE: GP-70660

; CURRENT APPLICATION NUMBER: US/09/713.158

; PRIOR FILING DATE: 2000-11-15

; PRIOR APPLICATION NUMBER: 60/165,800

; PRIOR FILING DATE: 1999-11-16

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 501

; TYPE: PRT

; ORGANISM: MUS MUSCULUS

9-713-158-2

Query Match 96.9%; Score 2582; DB 4; Length 501;
Best Local Similarity 96.6%; Pred. No. 1.4e-259;
Matches 484; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MAQALPMLLLMWGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Db 1 MAQALPMLLLMWGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Qy 61 VEMVDNLKRGSGQGYVEMTVGSPPTNLILVDTGSSNFVAVCAAPHPHLYRQQLSST 120
Db 61 VEMVDNLKRGSGQGYVEMTVGSPPTNLILVDTGSSNFVAVCAAPHPHLYRQQLSST 120
Qy 121 YRDLKRGVYPTQGWEGELGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLKRGVYPTQGWEGELGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLPPFDLSLVKQTHVNLPSLHLCGAGFPLNOSSEVLASVGGSMIIGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVNLPSLHLCGAGFPLNOSSEVLASVGGSMIIGI 240
Qy 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYVDSGTTNLRPKK 300
Db 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYVDSGTTNLRPKK 300

QY 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKIGFAVSAC 420
 DB 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKIGFAVSAC 420
 QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
 DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
 QY 481 RCLRLRQOHHDDFADDSILLK 501
 DB 481 RCLRLRQOHHDDFADDSILLK 501

RESULT 6
 US-09-548-367D-8
 ; Sequence 8, Application US/09548367D
 ; Patent No. 6440698
 ; GENERAL INFORMATION:
 ; APPLICANT: GURNEY ET AL.
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
 ; FILE REFERENCE: 29915/6280H
 ; CURRENT FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: US 09/548,367D
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: US 60/155,493
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: US 09/404,133
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: US 60/101,594
 ; PRIOR FILING DATE: 1998-09-24
 ; NUMBER OF SEQ ID NOS: 73
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-548-367D-8

Query Match 96.4%; Score 2567; DB 4; Length 501;
 Best Local Similarity 96.2%; Pred. No. 5.1e-258;
 Matches 482; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
 QY 1 MAQALPWLMLWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
 1 MAPALHLLWLLWVSGMLPAAQGTGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
 61 VEMVDNLKRGKSGQYVVTGSGPPQTLNVLVDGSSNFVGAAPHPFLHRYQRLSST 120
 61 VEMVDNLKRGKSGQYVVTGSGPPQTLNVLVDGSSNFVGAAPHPFLHRYQRLSST 120
 121 YRDLRKGVPVYPTQKWEGLGTDLVSPHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
 121 YRDLRKGVPVYPTQKWEGLGTDLVSPHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
 181 GLAYAEIARPDSDLEPFDDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
 181 GLAYAEIARPDSDLEPFDDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTGTTNLRPKK 300
 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTGTTNLRPKK 300
 301 VFEAAVKSIIKAASSTKFPDGFGLWGLVQWAGTTPWNIFFVISLYLMGEVNTNOSFRIT 360
 301 VFEAAVKSIIKAASSTKFPDGFGLWGLVQWAGTTPWNIFFVISLYLMGEVNTNOSFRIT 360
 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKIGFAVSAC 420
 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKIGFAVSAC 420

QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
 DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
 QY 481 RCLRLRQOHHDDFADDSILLK 501
 DB 481 RCLRLRQOHHDDFADDSILLK 501
 RESULT 7
 US-09-548-372D-6
 ; Sequence 6, Application US/09548372D
 ; Patent No. 6420534
 ; GENERAL INFORMATION:
 ; APPLICANT: GURNEY ET AL.
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
 ; FILE REFERENCE: 29915/6280I
 ; CURRENT FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: US 09/548,372D
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: US 60/155,493
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: US 09/404,133
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: US 60/101,594
 ; PRIOR FILING DATE: 1998-09-24
 ; NUMBER OF SEQ ID NOS: 73
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 476
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-548-372D-6

Query Match 94.1%; Score 2506.5; DB 4; Length 476;
 Best Local Similarity 95.0%; Pred. No. 9.1e-252;
 Matches 476; Conservative 0; Mismatches 0; Indels 25; Gaps 1;
 QY 1 MAQALPWLMLWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
 1 MAQALPWLMLWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
 61 VEMVDNLKRGKSGQYVVTGSGPPQTLNVLVDGSSNFVGAAPHPFLHRYQRLSST 120
 61 VEMVDNLKRGKSGQYVVTGSGPPQTLNVLVDGSSNFVGAAPHPFLHRYQRLSST 120
 121 YRDLRKGVPVYPTQKWEGLGTDLVSPHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
 121 YRDLRKGVPVYPTQKWEGLGTDLVSPHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
 181 GLAYAEIARPDSDLEPFDDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
 181 GLAYAEIARPDSDLEPFDDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTGTTNLRPKK 300
 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTGTTNLRPKK 300
 301 VFEAAVKSIIKAASSTKFPDGFGLWGLVQWAGTTPWNIFFVISLYLMGEVNTNOSFRIT 360
 301 VFEAAVKSIIKAASSTKFPDGFGLWGLVQWAGTTPWNIFFVISLYLMGEVNTNOSFRIT 360
 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKIGFAVSAC 420
 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKIGFAVSAC 420
 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
 481 RCLRLRQOHHDDFADDSILLK 501

Db 456 RCLRCLRQHQHDDFADDISLLK 476

RESULT 8

US-09-548-367D-6
; Sequence 6, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548.367D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-6

Query Match 94.1%; Score 2506.5; DB 4; Length 476;
Best Local Similarity 95.0%; Pred. No. 9.1e-252;
Matches 476; Conservative 0; Mismatches 0; Indels 25; Gaps 1;
Qy 1 MAQALPMLLLMMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Db 1 MAQALPMLLLMMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Qy 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNLIIVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNLIIVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
Qy 121 YRDLKRGVYVPTGKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLKRGVYVPTGKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMITGGI 240
Db 181 GLAYAEIARPDSDLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMITGGI 215
Qy 241 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTTNLRLPKK 300
Db 216 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTTNLRLPKK 275
Qy 301 VFEAAVKSIAASSTKFPDGFGLGEOLVVCWAGTTPWNIIPVISLXLMGEVTVNQSPRIT 360
Db 276 VFEAAVKSIAASSTKFPDGFGLGEOLVVCWAGTTPWNIIPVISLXLMGEVTVNQSPRIT 335
Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGYVYVDFRARRKRGFAVSAC 420
Db 336 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGYVYVDFRARRKRGFAVSAC 395
Qy 421 HVHDEFRTAAGVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
Db 396 HVHDEFRTAAGVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 455
Qy 481 RCLRCLRQHQHDDFADDISLLK 501
Db 456 RCLRCLRQHQHDDFADDISLLK 476

RESULT 9

US-09-548-372D-73
; Sequence 73, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548.372D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 73
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-548-372D-73

Query Match 90.9%; Score 2420.5; DB 4; Length 476;
Best Local Similarity 91.8%; Pred. No. 8e-243;
Matches 460; Conservative 5; Mismatches 11; Indels 25; Gaps 1;
Qy 1 MAQALPMLLLMMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Db 1 MAQALPMLLLMMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Qy 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNLIIVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNLIIVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
Qy 121 YRDLKRGVYVPTGKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLKRGVYVPTGKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMITGGI 240
Db 181 GLAYAEIARPDSDLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMITGGI 215
Qy 241 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTTNLRLPKK 300
Db 216 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTTNLRLPKK 275
Qy 301 VFEAAVKSIAASSTKFPDGFGLGEOLVVCWAGTTPWNIIPVISLXLMGEVTVNQSPRIT 360
Db 276 VFEAAVKSIAASSTKFPDGFGLGEOLVVCWAGTTPWNIIPVISLXLMGEVTVNQSPRIT 335
Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGYVYVDFRARRKRGFAVSAC 420
Db 336 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGYVYVDFRARRKRGFAVSAC 395
Qy 421 HVHDEFRTAAGVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
Db 396 HVHDEFRTAAGVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 455
Qy 481 RCLRCLRQHQHDDFADDISLLK 501
Db 456 RCLRCLRQHQHDDFADDISLLK 476

RESULT 10

US-09-548-367D-73
; Sequence 73, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 73
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
; US-548-367D-73
Query Match
Best Local Similarity 90.9%; Score 2420.5; DB 4; Length 476;
Matches 460; Conservative 5; Mismatches 11; Indels 25; Gaps 1;
QY 1 MAQALPWLWAGVLPAGHTQHGIPLRLSLGGLGAPLGLRLPRETDEPEEPRGRGSF 60
DB 1 MAPALHLLWVSGMLPAQHTLGIPLRLSLGGLGAPLGLRLPRETDEEPEEPRGRGSF 60
QY 61 VEMVDNLKSGGGYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQRLSST 120
DB 61 VEMVDNLKSGGGYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQRLSST 120
QY 121 YRDLRGVVPYTOGKWEGLGTDLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
DB 121 YRDLRGVVPYTOGKWEGLGTDLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLEPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
DB 181 GLAYAEIARPDSDLEPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
QY 241 DLSLYTGSLSWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
DB 241 DLSLYTGSLSWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
QY 216 DLSLYTGSLSWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 275
DB 216 DLSLYTGSLSWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 275
QY 301 VFEAAVKSIAASTKTEKFPDGLGOLVCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 360
DB 301 VFEAAVKSIAASTKTEKFPDGLGOLVCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 360
QY 276 VFEAAVKSIAASTKTEKFPDGLGOLVCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 335
DB 276 VFEAAVKSIAASTKTEKFPDGLGOLVCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 335
QY 361 ILPOQYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420
DB 361 ILPOQYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420
QY 336 ILPOQYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 395
DB 336 ILPOQYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 395
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALEMPLCLMVCOW 480
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALEMPLCLMVCOW 480
QY 396 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALEMPLCLMVCOW 455
DB 396 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALEMPLCLMVCOW 455
QY 481 RCLRLCRLQHQHDDFADDISLLK 501
DB 481 RCLRLCRLQHQHDDFADDISLLK 476

RESULT 11
US-09-548-372D-30
; Sequence 30, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 30
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-548-372D-30
Query Match
Best Local Similarity 99.8%; Pred. No. 2e-240;
Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAQALPWLWAGVLPAGHTQHGIPLRLSLGGLGAPLGLRLPRETDEPEEPRGRGSF 60
DB 1 MAQALPWLWAGVLPAGHTQHGIPLRLSLGGLGAPLGLRLPRETDEPEEPRGRGSF 60
QY 61 VEMVDNLKSGGGYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQRLSST 120
DB 61 VEMVDNLKSGGGYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQRLSST 120
QY 121 YRDLRGVVPYTOGKWEGLGTDLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
DB 121 YRDLRGVVPYTOGKWEGLGTDLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLEPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
DB 181 GLAYAEIARPDSDLEPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
QY 241 DLSLYTGSLSWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
DB 241 DLSLYTGSLSWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
QY 301 VFEAAVKSIAASTKTEKFPDGLGOLVCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 360
DB 301 VFEAAVKSIAASTKTEKFPDGLGOLVCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 360
QY 361 ILPOQYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420
DB 361 ILPOQYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALEMPLCLMVCOW 453
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALEMPLCLMVCOW 453

RESULT 12
US-09-548-367D-30
; Sequence 30, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 30

; LENGTH: 453

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-548-367D-30

Query Match 90.0%; Score 2397; DB 4; Length 453;
Best Local Similarity 99.8%; Pred. No. 2e-240;
Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 MAQALPWLMLWMGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLMLWMGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Qy 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNIIIVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNIIIVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Qy 121 YRDLRKGYVYPYTOGKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLRKGYVYPYTOGKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEATARPDDSLPEPFDLSLVKQTHVFNLSHLGAGFPLNOSSEVLASVGGSMIIGGI 240
Db 181 GLAYAEATARPDDSLPEPFDLSLVKQTHVFNLSHLGAGFPLNOSSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSCTTNLRPKK 300
Db 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSCTTNLRPKK 300
Qy 301 VFEAAVKSIIKAASSTKFPDGFGLGEQLVCWQAGTTPWNIPFVLSLYLMGEVTTNQSFRIT 360
Db 301 VFEAAVKSIIKAASSTKFPDGFGLGEQLVCWQAGTTPWNIPFVLSLYLMGEVTTNQSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDDCYKFAISOSSTCTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDDCYKFAISOSSTCTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
Qy 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDES 453
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDES 453
```

RESULT 13

US-09-548-372D-32

; Sequence 32, Application US/09548372D

; Patent No. 6420334

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; FILE REFERENCE: 29915/62801

; CURRENT APPLICATION NUMBER: US/09/548,372D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 32

; LENGTH: 459

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-548-372D-32

Query Match 90.0%; Score 2397; DB 4; Length 459;

Best Local Similarity 99.8%; Pred. No. 2.1e-240;

Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MAQALPWLMLWMGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLMLWMGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Qy 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNIIIVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNIIIVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Qy 121 YRDLRKGYVYPYTOGKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLRKGYVYPYTOGKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEATARPDDSLPEPFDLSLVKQTHVFNLSHLGAGFPLNOSSEVLASVGGSMIIGGI 240
Db 181 GLAYAEATARPDDSLPEPFDLSLVKQTHVFNLSHLGAGFPLNOSSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSCTTNLRPKK 300
Db 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSCTTNLRPKK 300
Qy 301 VFEAAVKSIIKAASSTKFPDGFGLGEQLVCWQAGTTPWNIPFVLSLYLMGEVTTNQSFRIT 360
Db 301 VFEAAVKSIIKAASSTKFPDGFGLGEQLVCWQAGTTPWNIPFVLSLYLMGEVTTNQSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDDCYKFAISOSSTCTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDDCYKFAISOSSTCTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
Qy 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDES 453
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDES 453
```

RESULT 14

US-09-548-367D-32

; Sequence 32, Application US/09548367D

; Patent No. 6440698

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE

; FILE REFERENCE: 29915/6280H

; CURRENT APPLICATION NUMBER: US/09/548,367D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 32

; LENGTH: 459

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-548-367D-32

Query Match 90.0%; Score 2397; DB 4; Length 459;

Best Local Similarity 99.8%; Pred. No. 2.1e-240;

Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 MAQALPWLMLWMGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLMLWMGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Qy 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNIIIVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPQTLNIIIVDTGSSNFVAGAAAPHPLHRYQRLSST 120
```

Query Match 86.9%; Score 2315; DB 4; Length 774;
Best Local Similarity 98.6%; Pred. No. 1.7e-231;
Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 58 GSFVEMVDNLKRGSGQYVEMTVGSPPTLNILVDTGSSNFAYGAAPHPFLHRYORQL 117
DB 1 GSFVEMVDNLKRGSGQYVEMTVGSPPTLNILVDTGSSNFAYGAAPHPFLHRYORQL 60
QY 118 SSTYRDLKRGVYVPTQKWEGLGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWE 177
DB 61 SSTYRDLKRGVYVPTQKWEGLGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWE 120
QY 178 GTGLAYAEIARPDSDLEPFFDSLKVQTHVNLFSHLGAGFPLNQSEVLASVGGSMII 237
DB 121 GTGLAYAEIARPDSDLEPFFDSLKVQTHVNLFSHLGAGFPLNQSEVLASVGGSMII 180
QY 238 GGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIIVDSGTTNLR 297
DB 181 GGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIIVDSGTTNLR 240
QY 298 PKKVFEEAAVKSIIKAASSTKTEKFPDGFMLGEOQVLCWQAGTTPWNIIPVVISLYLMGEV 357
DB 241 PKKVFEEAAVKSIIKAASSTKTEKFPDGFMLGEOQVLCWQAGTTPWNIIPVVISLYLMGEV 300
QY 358 RTILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKRIGFAV 417
DB 301 RTILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKRIGFAV 360
QY 418 SACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALEMPLCLMV 477
DB 361 SACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALEMPLCLMV 420
QY 478 CQWRCLRLCQHQHDDFADDISLLK 501
DB 421 CQWRCLRLCQHQHDDFADDISLLK 444

Search completed: June 5, 2003, 16:18:45
Job time : 29 secs

QY 121 YRDLKRGVYVPTQKWEGLGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWEGL 180
DB 121 YRDLKRGVYVPTQKWEGLGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWEGL 180
QY 181 GLAYAEIARPDSDLEPFFDSLKVQTHVNLFSHLGAGFPLNQSEVLASVGGSMIIIGI 240
DB 181 GLAYAEIARPDSDLEPFFDSLKVQTHVNLFSHLGAGFPLNQSEVLASVGGSMIIIGI 240
QY 241 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKK 300
DB 241 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKK 300
QY 301 VFEAAVKSIIKAASSTKTEKFPDGFMLGEOQVLCWQAGTTPWNIIPVVISLYLMGEV 360
DB 301 VFEAAVKSIIKAASSTKTEKFPDGFMLGEOQVLCWQAGTTPWNIIPVVISLYLMGEV 360
QY 361 ILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKRIGFAVSAC 420
DB 361 ILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKRIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDES 453
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDES 453

RESULT 15
US-09-009-191-4
Sequence 4, Application US/09009191
Patent No. 6319689
GENERAL INFORMATION:
APPLICANT: POWELL, DAVID
APPLICANT: CHAPMAN, CONRAD
APPLICANT: MURPHY, KAY
APPLICANT: SMITH, TRUDI
TITLE OF INVENTION: ASP2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,191
FILING DATE: 20-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9701684.4
FILING DATE: 28-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70368
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-191-4

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:12:20 ; Search time 24 Seconds
(without alignments)
2155.141 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLLMGAGVLPAAH.....CLRLRQHHDFADDISLLK 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

1 number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2564	100.0	501	9	US-09-869-414-4
2	2664	100.0	501	10	US-09-794-927-4
3	2664	100.0	501	10	US-09-795-847-4
4	2664	100.0	501	10	US-09-794-743-4
5	2664	100.0	501	10	US-09-794-748-4
6	2664	100.0	501	10	US-09-794-925-4
7	2664	100.0	501	10	US-09-681-442-4
8	2656	99.7	501	9	US-10-032-818-4
9	2656	99.7	501	9	US-10-214-932-104
10	2650	99.5	501	9	US-09-969-671A-2
11	2582	96.9	488	9	US-09-795-903A-2
12	2582	96.9	488	9	US-10-032-818-2
13	2582	96.9	488	10	US-09-796-264-2
14	2582	96.9	488	10	US-09-845-226-2
15	2582	96.9	503	9	US-09-795-903A-3
16	2582	96.9	503	9	US-10-032-818-3
17	2582	96.9	503	10	US-09-796-264-3
18	2582	96.9	503	10	US-09-845-226-3
19	2567	96.4	501	9	US-09-869-414-8

20	2567	96.4	501	10	US-09-794-927-8	Sequence 8, Appli
21	2567	96.4	501	10	US-09-795-847-8	Sequence 8, Appli
22	2567	96.4	501	10	US-09-794-743-8	Sequence 8, Appli
23	2567	96.4	501	10	US-09-794-748-8	Sequence 8, Appli
24	2567	96.4	501	10	US-09-794-925-8	Sequence 8, Appli
25	2567	96.4	501	10	US-09-681-442-8	Sequence 8, Appli
26	2506.5	94.1	476	9	US-09-869-414-6	Sequence 6, Appli
27	2506.5	94.1	476	10	US-09-794-927-6	Sequence 6, Appli
28	2506.5	94.1	476	10	US-09-795-847-6	Sequence 6, Appli
29	2506.5	94.1	476	10	US-09-794-743-6	Sequence 6, Appli
30	2506.5	94.1	476	10	US-09-794-748-6	Sequence 6, Appli
31	2506.5	94.1	476	10	US-09-794-925-6	Sequence 6, Appli
32	2506.5	94.1	476	10	US-09-681-442-6	Sequence 6, Appli
33	2397	90.0	453	9	US-09-869-414-30	Sequence 30, Appl
34	2397	90.0	453	10	US-09-794-927-30	Sequence 30, Appl
35	2397	90.0	453	10	US-09-795-847-30	Sequence 30, Appl
36	2397	90.0	453	10	US-09-794-743-30	Sequence 30, Appl
37	2397	90.0	453	10	US-09-794-748-30	Sequence 30, Appl
38	2397	90.0	453	10	US-09-794-925-30	Sequence 30, Appl
39	2397	90.0	453	10	US-09-681-442-30	Sequence 30, Appl
40	2397	90.0	459	9	US-09-869-414-32	Sequence 32, Appl
41	2397	90.0	459	10	US-09-794-927-32	Sequence 32, Appl
42	2397	90.0	459	10	US-09-795-847-32	Sequence 32, Appl
43	2397	90.0	459	10	US-09-794-743-32	Sequence 32, Appl
44	2397	90.0	459	10	US-09-794-748-32	Sequence 32, Appl
45	2397	90.0	459	10	US-09-794-925-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1

US-09-869-414-4
; Sequence 4, Application US/09869414
; Publication No. US20030077226A1
; GENERAL INFORMATION:
; APPLICANT: Beinkowski et al.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND US
; FILE REFERENCE: 28341/6280M
; CURRENT APPLICATION NUMBER: US/09/869,414
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-869-414-4

Query Match 100.0%; Score 2664; DB 9; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAQALPWLMLLMGAGVLPAAHGTQHGIRLPLRSLGGLGAPLGLRLPLRETDEEPEEPCRRGSF	60
DB	1	MAQALPWLMLLMGAGVLPAAHGTQHGIRLPLRSLGGLGAPLGLRLPLRETDEEPEEPCRRGSF	60
QY	61	VEVDNLKRGSCQGYVEMTVGSPQTNLILVDTCSSNFVAGCAAPHLPHRYQRLSST	120
DB	61	VEVDNLKRGSCQGYVEMTVGSPQTNLILVDTCSSNFVAGCAAPHLPHRYQRLSST	120
QY	121	YRDLKRGVIVPYTQCKWEGELGTDLVSIPHGPNVTVRANIAITEDSKFFINGSNWEGIL	180

Db 121 YRDURKGVYVYPTQCKWEGELGTDLSVPHGPNVTVRANIAAITESSDKFFINGSNWEGIL 180
181 GLAYAEIARPDSDLEPFDFSLVKQTHVFNLFSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
181 GLAYAEIARPDSDLEPFDFSLVKQTHVFNLFSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
241 DHSLYTGSLSWTPTRRWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDVSSTTNLRPKK 300
241 DHSLYTGSLSWTPTRRWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDVSSTTNLRPKK 300
301 VFEAAVKSIAKASSTERKFPDGFNLCEQLVQWAGTTPWNIIPVVISLYLMGEVTVNQSFRT 360
301 VFEAAVKSIAKASSTERKFPDGFNLCEQLVQWAGTTPWNIIPVVISLYLMGEVTVNQSFRT 360
361 ILPOQYLRPEVDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIGFAVSAC 420
361 ILPOQYLRPEVDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIGFAVSAC 420
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMWQW 480
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMWQW 480
481 RCLRCLROQHDDFADDISLLK 501
481 RCLRCLROQHDDFADDISLLK 501

RESULT 2
US-09-794-927-4
; Sequence 4, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794,927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-927-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQALPWLWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGSF 60
Db 1 MAQALPWLWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGSF 60
QY 61 VEMVDNLRKSGQGYVEMTVGSPPTNLINLVDTSNFAVGAAPHFLHRYQRLSST 120
Db 61 VEMVDNLRKSGQGYVEMTVGSPPTNLINLVDTSNFAVGAAPHFLHRYQRLSST 120

QY 121 YRDURKGVYVYPTQCKWEGELGTDLSVPHGPNVTVRANIAAITESSDKFFINGSNWEGIL 180
Db 121 YRDURKGVYVYPTQCKWEGELGTDLSVPHGPNVTVRANIAAITESSDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLEPFDFSLVKQTHVFNLFSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDLEPFDFSLVKQTHVFNLFSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
QY 241 DHSLYTGSLSWTPTRRWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDVSSTTNLRPKK 300
Db 241 DHSLYTGSLSWTPTRRWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDVSSTTNLRPKK 300
QY 301 VFEAAVKSIAKASSTERKFPDGFNLCEQLVQWAGTTPWNIIPVVISLYLMGEVTVNQSFRT 360
Db 301 VFEAAVKSIAKASSTERKFPDGFNLCEQLVQWAGTTPWNIIPVVISLYLMGEVTVNQSFRT 360
QY 361 ILPOQYLRPEVDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIGFAVSAC 420
Db 361 ILPOQYLRPEVDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMWQW 480
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMWQW 480
QY 481 RCLRCLROQHDDFADDISLLK 501
Db 481 RCLRCLROQHDDFADDISLLK 501

RESULT 3
US-09-795-847-4
; Sequence 4, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280DE
; CURRENT APPLICATION NUMBER: US/09/795,847
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-847-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQALPWLWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGSF 60
Db 1 MAQALPWLWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGSF 60
QY 61 VEMVDNLRKSGQGYVEMTVGSPPTNLINLVDTSNFAVGAAPHFLHRYQRLSST 120

Db 61 VEMVDNLGRSGGQYVEMTVGSPQTLNILDVTGSSNFVGAAPHFLHRYQRLSST 120
Qy 121 YRDLRKGVVYPYTOCKWEGELGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLRKGVVYPYTOCKWEGELGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDSLEPFDDSLVKQTHVPNLSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDSLEPFDDSLVKQTHVPNLSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGSWLYTPIRREWEYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
Db 241 DHSLYTGSWLYTPIRREWEYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
Qy 301 VFEAAVKSIAAASSTERKFPDGFGLGEOLVCMQAGTTTWNIPFPVLSLYLMGEVTVNOSFRIT 360
Db 301 VFEAAVKSIAAASSTERKFPDGFGLGEOLVCMQAGTTTWNIPFPVLSLYLMGEVTVNOSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGFYVVDRAKRKRGFAVSAC 420
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQW 480
Qy 481 RCLRCLRQHQHDFADDDISLLK 501
Db 481 RCLRCLRQHQHDFADDDISLLK 501

RESULT 4
US-09-794-743-4
; Sequence 4, Application US/09794743
; Patent No. US2001021391A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280BC
; CURRENT APPLICATION NUMBER: US/09/794,743
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-743-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAQALPWLMMGAGVLPFAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLMMGAGVLPFAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60

Qy 61 VEMVDNLGRSGGQYVEMTVGSPQTLNILDVTGSSNFVGAAPHFLHRYQRLSST 120
Db 61 VEMVDNLGRSGGQYVEMTVGSPQTLNILDVTGSSNFVGAAPHFLHRYQRLSST 120
Qy 121 YRDLRKGVVYPYTOCKWEGELGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLRKGVVYPYTOCKWEGELGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDSLEPFDDSLVKQTHVPNLSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDSLEPFDDSLVKQTHVPNLSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGSWLYTPIRREWEYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
Db 241 DHSLYTGSWLYTPIRREWEYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
Qy 301 VFEAAVKSIAAASSTERKFPDGFGLGEOLVCMQAGTTTWNIPFPVLSLYLMGEVTVNOSFRIT 360
Db 301 VFEAAVKSIAAASSTERKFPDGFGLGEOLVCMQAGTTTWNIPFPVLSLYLMGEVTVNOSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGFYVVDRAKRKRGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGFYVVDRAKRKRGFAVSAC 420
Qy 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQW 480
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQW 480
Qy 481 RCLRCLRQHQHDFADDDISLLK 501
Db 481 RCLRCLRQHQHDFADDDISLLK 501

RESULT 5
US-09-794-748-4
; Sequence 4, Application US/09794748
; Patent No. US2002003731A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280JL
; CURRENT APPLICATION NUMBER: US/09/794,748
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-748-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAQALPWLMMGAGVLPFAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60

Db 1 MAQALPWL... 60
Qy 61 VEMVDNL... 120
Db 61 VEMVDNL... 120
Qy 121 YRDLRKG... 180
Db 121 YRDLRKG... 180
Qy 181 GLAYAEI... 240
Db 181 GLAYAEI... 240
Qy 241 DHSLYT... 300
Db 241 DHSLYT... 300
Qy 301 VFEEAV... 360
Db 301 VFEEAV... 360
Qy 361 ILPQOY... 420
Db 361 ILPQOY... 420
Qy 421 HVHDEF... 480
Db 421 HVHDEF... 480
Qy 481 RCLRL... 501
Db 481 RCLRL... 501

RESULT 6
US-09-794-925-4
; Sequence 4, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR
; FILE REFERENCE: 28341/6280H1
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-925-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQALPWL... 60
Db 1 MAQALPWL... 60
Qy 61 VEMVDNL... 120
Db 61 VEMVDNL... 120
Qy 121 YRDLRKG... 180
Db 121 YRDLRKG... 180
Qy 181 GLAYAEI... 240
Db 181 GLAYAEI... 240
Qy 241 DHSLYT... 300
Db 241 DHSLYT... 300
Qy 301 VFEEAV... 360
Db 301 VFEEAV... 360
Qy 361 ILPQOY... 420
Db 361 ILPQOY... 420
Qy 421 HVHDEF... 480
Db 421 HVHDEF... 480
Qy 481 RCLRL... 501
Db 481 RCLRL... 501

RESULT 7
US-09-681-442-4
; Sequence 4, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT FILING DATE: 2001-04-05
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-681-442-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSF 60
Dbs 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSF 60
QY 61 VEMVDNLKCKSGQGYVEMTVGSPPTNLILVDTCSSNFAYGAAAPHPFLHRYQRLSST 120
Dbs 61 VEMVDNLKCKSGQGYVEMTVGSPPTNLILVDTCSSNFAYGAAAPHPFLHRYQRLSST 120
QY 121 YRDLRKGVPYPTQCKWEGELGTDLSVIPHGNVTVRANIAAITSDEKFFINGSNWEGIL 180
Dbs 121 YRDLRKGVPYPTQCKWEGELGTDLSVIPHGNVTVRANIAAITSDEKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSHLGCGAGFPLNQSEVLASVGGSMIIGI 240
Dbs 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSHLGCGAGFPLNQSEVLASVGGSMIIGI 240
QY 241 DHSLYTGSWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVGTTNLRPKK 300
Dbs 241 DHSLYTGSWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVGTTNLRPKK 300
QY 301 VFEAAVKSIAASSTKEKFPDGFMLGEOQLVCWQAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
Dbs 301 VFEAAVKSIAASSTKEKFPDGFMLGEOQLVCWQAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
QY 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFDRARKRIGFAVSAC 420
Dbs 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFDRARKRIGFAVSAC 420
QY 421 HVHDEFRTAAVGEFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
Dbs 421 HVHDEFRTAAVGEFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
QY 481 RCLRLRQOHHDFADDISLLK 501
Dbs 481 RCLRLRQOHHDFADDISLLK 501

RESULT 8

US-10-032-818-4
; Sequence 4, Application US/10032818
; Publication No. US20030092629A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Koelsch, Gerald
; APPLICANT: Ghosh, Arun K.
; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
; FILE REFERENCE: 2932 1006-007
; CURRENT APPLICATION NUMBER: US/10/032,818
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/275,756
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/258,705
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-032-818-4

Query Match 99.7%; Score 2656; DB 9; Length 501;
Best Local Similarity 99.8%; Pred. No. 2.9e-240;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSF 60
Dbs 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSF 60
QY 61 VEMVDNLKCKSGQGYVEMTVGSPPTNLILVDTCSSNFAYGAAAPHPFLHRYQRLSST 120
Dbs 61 VEMVDNLKCKSGQGYVEMTVGSPPTNLILVDTCSSNFAYGAAAPHPFLHRYQRLSST 120

QY 121 YRDLRKGVPYPTQCKWEGELGTDLSVIPHGNVTVRANIAAITSDEKFFINGSNWEGIL 180
Dbs 121 YRDLRKGVPYPTQCKWEGELGTDLSVIPHGNVTVRANIAAITSDEKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSHLGCGAGFPLNQSEVLASVGGSMIIGI 240
Dbs 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSHLGCGAGFPLNQSEVLASVGGSMIIGI 240
QY 241 DHSLYTGSWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVGTTNLRPKK 300
Dbs 241 DHSLYTGSWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVGTTNLRPKK 300
QY 301 VFEAAVKSIAASSTKEKFPDGFMLGEOQLVCWQAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
Dbs 301 VFEAAVKSIAASSTKEKFPDGFMLGEOQLVCWQAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
QY 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFDRARKRIGFAVSAC 420
Dbs 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFDRARKRIGFAVSAC 420
QY 421 HVHDEFRTAAVGEFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
Dbs 421 HVHDEFRTAAVGEFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
QY 481 RCLRLRQOHHDFADDISLLK 501
Dbs 481 RCLRLRQOHHDFADDISLLK 501

RESULT 9

US-10-214-932-104
; Sequence 104, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhan
; APPLICANT: KIM, Dae Heon
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: APB02/US
; CURRENT APPLICATION NUMBER: US/10/214,932
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 104
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-214-932-104

Query Match 99.7%; Score 2656; DB 9; Length 501;
Best Local Similarity 99.8%; Pred. No. 2.9e-240;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSF 60
Dbs 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSF 60
QY 61 VEMVDNLKCKSGQGYVEMTVGSPPTNLILVDTCSSNFAYGAAAPHPFLHRYQRLSST 120
Dbs 61 VEMVDNLKCKSGQGYVEMTVGSPPTNLILVDTCSSNFAYGAAAPHPFLHRYQRLSST 120
QY 121 YRDLRKGVPYPTQCKWEGELGTDLSVIPHGNVTVRANIAAITSDEKFFINGSNWEGIL 180
Dbs 121 YRDLRKGVPYPTQCKWEGELGTDLSVIPHGNVTVRANIAAITSDEKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSHLGCGAGFPLNQSEVLASVGGSMIIGI 240
Dbs 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSHLGCGAGFPLNQSEVLASVGGSMIIGI 240
QY 241 DHSLYTGSWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVGTTNLRPKK 300
Dbs 241 DHSLYTGSWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVGTTNLRPKK 300

361 ILPQOYLRPVEDVATSDDCYKFAISQSGTGMGAVIMEGFYVVDNRKRIGFAVSAC 420
 361 ILPQOYLRPVEDVATSDDCYKFAISQSGTGMGAVIMEGFYVVDNRKRIGFAVSAC 420
 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
 481 RCLRLRQOHHDFADDISLLK 501
 481 RCLRLRQOHHDFADDISLLK 501

RESULT 11

US-09-795-903A-2
 ; Sequence 2, Application US/09795903A
 ; Patent No. US20020164760A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Jordan J.N.
 ; APPLICANT: Koelsch, Gerald
 ; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
 ; FILE REFERENCE: OMRF 179
 ; CURRENT APPLICATION NUMBER: US/09795.903A
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 09/604,608
 ; PRIOR FILING DATE: 2000-06-27
 ; PRIOR APPLICATION NUMBER: 60/168,060
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: 60/177,836
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR APPLICATION NUMBER: 60/178,368
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: 60/210,292
 ; PRIOR FILING DATE: 2000-06-08
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: Patent Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Purified Memapsin 2
 ; OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide
 ; OTHER INFORMATION: residues:
 ; OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,
 ; OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 281, 283, and
 ; OTHER INFORMATION: 376-377 are residues in contact with the OM99-2
 ; OTHER INFORMATION: inhibitor
 ; OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,
 ; OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and
 ; OTHER INFORMATION: 220-224 are N-lobe Beta Strands
 ; OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices
 ; OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,
 ; OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-349,
 ; OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,
 ; OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands
 ; OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,
 ; OTHER INFORMATION: and 427-431 are C-lobe Helices
 ; US-09-795-903A-2

Query Match 96.9%; Score 2582; DB 9; Length 488;
 Best Local Similarity 99.8%; Pred. No. 2.4e-233;
 Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 14 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSGFVMDNLRKSGQ 73
 Db 1 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSGFVMDNLRKSGQ 60
 Qy 74 GYVYEMTVGSPPTNLILVDGTSSNFVAGAAPHPFLHRYQRLSSTYRDLRGVVPYT 133
 Db 61 GYVYEMTVGSPPTNLILVDGTSSNFVAGAAPHPFLHRYQRLSSTYRDLRGVVPYT 120

301 VFEAAVKSIIKAASSTKPPDGFMLGEOLVCWAGTTPWNIPIVSLYLMGEVNTQSFRT 360
 301 VFEAAVKSIIKAASSTKPPDGFMLGEOLVCWAGTTPWNIPIVSLYLMGEVNTQSFRT 360
 361 ILPQOYLRPVEDVATSDDCYKFAISQSGTGMGAVIMEGFYVVDNRKRIGFAVSAC 420
 361 ILPQOYLRPVEDVATSDDCYKFAISQSGTGMGAVIMEGFYVVDNRKRIGFAVSAC 420
 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
 481 RCLRLRQOHHDFADDISLLK 501
 481 RCLRLRQOHHDFADDISLLK 501

RESULT 10

US-09-969-671A-2
 ; Sequence 2, Application US/09969671A
 ; Publication No. US20030036112A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHAPMAN, CONRAD G.
 ; APPLICANT: MURPHY, KAY
 ; APPLICANT: POWELL, DAVID J.
 ; APPLICANT: SMITH, TRUDI S.
 ; TITLE OF INVENTION: ASP2
 ; FILE REFERENCE: GH-70368-D1
 ; CURRENT APPLICATION NUMBER: US/09/969,671A
 ; CURRENT FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: UK 9701684.4
 ; PRIOR FILING DATE: 1997-01-28
 ; PRIOR APPLICATION NUMBER: 09/009,191
 ; PRIOR FILING DATE: 1998-01-20
 ; PRIOR APPLICATION NUMBER: 09/694,200
 ; PRIOR FILING DATE: 2000-10-23
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-969-671A-2

Query Match 99.5%; Score 2650; DB 9; Length 501;
 Best Local Similarity 99.6%; Pred. No. 1.1e-239;
 Matches 499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MAQALPWLMLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSGF 60
 1 MAQALPWLMLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSGF 60
 61 VEMVDNLRKSGQGYVEMTVGSPPTNLILVDGTSSNFVAGAAPHPFLHRYQRLSST 120
 61 VEMVDNLRKSGQGYVEMTVGSPPTNLILVDGTSSNFVAGAAPHPFLHRYQRLSST 120
 121 YRDLRGVVPYTOCKWEGELGTLVSIPIHGPNTVTRANIAAITESSDKFFINGSNWEGIL 180
 121 YRDLRGVVPYTOCKWEGELGTLVSIPIHGPNTVTRANIAAITESSDKFFINGSNWEGIL 180
 181 GLAYEAIARPDDSLPEFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIGGI 240
 181 GLAYEAIARPDDSLPEFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIGGI 240
 241 DSHLYTGSLSWYTPIRREWEYVILVVEINGQDLKMDCKEYNDKSIYDSGTNLRPKK 300
 241 DSHLYTGSLSWYTPIRREWEYVILVVEINGQDLKMDCKEYNDKSIYDSGTNLRPKK 300
 301 VFEAAVKSIIKAASSTKPPDGFMLGEOLVCWAGTTPWNIPIVSLYLMGEVNTQSFRT 360
 301 VFEAAVKSIIKAASSTKPPDGFMLGEOLVCWAGTTPWNIPIVSLYLMGEVNTQSFRT 360

QY 134 QKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDPS 193
 DB 121 QKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDPS 180
 QY 194 LEPPFDSLVKQTHVPNLFSLHLCAGFPLNQLVSEVLASVGGSMIIGGIDHSLYTSGLWYTP 253
 DB 181 LEPPFDSLVKQTHVPNLFSLHLCAGFPLNQLVSEVLASVGGSMIIGGIDHSLYTSGLWYTP 240
 QY 254 IRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTGTTNLRPLPKVFEAAVKSIRKAS 313
 DB 241 IRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTGTTNLRPLPKVFEAAVKSIRKAS 300
 QY 314 STEKPPDGFVLGEOQLVCWQAGTTPWNIPFVLSYLMGEVNTNOSFRITILPOOYLRLPVEDV 373
 DB 301 STEKPPDGFVLGEOQLVCWQAGTTPWNIPFVLSYLMGEVNTNOSFRITILPOOYLRLPVEDV 360
 QY 374 ATSDODCYKFAISOSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEG 433
 DB 361 ATSDODCYKFAISOSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEG 420
 QY 434 PFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVCMQWCLRLRQHQHDDF 493
 DB 421 PFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVCMQWCLRLRQHQHDDF 480
 QY 494 ADDISLLK 501
 DB 481 ADDISLLK 488

RESULT 12

US-10-032-818-2
 ; Sequence 2, Application US/10032818
 ; Publication No. US20030092629A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Jordan J.N.
 ; APPLICANT: Koelsch, Gerald
 ; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
 ; FILE REFERENCE: 2932.1006-007
 ; CURRENT APPLICATION NUMBER: US/10/032,818
 ; CURRENT FILING DATE: 2001-12-28
 ; PRIOR FILING DATE: 2001-12-28
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: US 60/258,705
 ; PRIOR FILING DATE: 2000-12-28
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-032-818-2

Query Match 96.9%; Score 2582; DB 9; Length 488;
 Best Local Similarity 99.8%; Pred. No. 2.4e-233;
 Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 14 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPRGRGSFVEMVDNLGRKSGQ 73
 DB 1 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPRGRGSFVEMVDNLGRKSGQ 60
 QY 74 GYVEMTVGSPQTLNVLVDGSSNFVAGAAPHPFLHRYQRLSSYVRLKRGVYVPYT 133
 DB 61 GYVEMTVGSPQTLNVLVDGSSNFVAGAAPHPFLHRYQRLSSYVRLKRGVYVPYT 120
 QY 134 QKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDPS 193
 DB 121 QKWEGLGTLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDPS 180
 QY 194 LEPPFDSLVKQTHVPNLFSLHLCAGFPLNQLVSEVLASVGGSMIIGGIDHSLYTSGLWYTP 253
 DB 181 LEPPFDSLVKQTHVPNLFSLHLCAGFPLNQLVSEVLASVGGSMIIGGIDHSLYTSGLWYTP 240

QY 254 IRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTGTTNLRPLPKVFEAAVKSIRKAS 313
 DB 241 IRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTGTTNLRPLPKVFEAAVKSIRKAS 300
 QY 314 STEKPPDGFVLGEOQLVCWQAGTTPWNIPFVLSYLMGEVNTNOSFRITILPOOYLRLPVEDV 373
 DB 301 STEKPPDGFVLGEOQLVCWQAGTTPWNIPFVLSYLMGEVNTNOSFRITILPOOYLRLPVEDV 360
 QY 374 ATSDODCYKFAISOSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEG 433
 DB 361 ATSDODCYKFAISOSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEG 420
 QY 434 PFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVCMQWCLRLRQHQHDDF 493
 DB 421 PFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVCMQWCLRLRQHQHDDF 480
 QY 494 ADDISLLK 501
 DB 481 ADDISLLK 488

RESULT 13

US-09-796-264-2
 ; Sequence 2, Application US/09796264
 ; Patent No. US20020049303A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Jordan J.N.
 ; APPLICANT: Lin, Xinli
 ; APPLICANT: Koelsch, Gerald
 ; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
 ; FILE REFERENCE: OMRF 179
 ; CURRENT APPLICATION NUMBER: US/09/796,264
 ; CURRENT FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 09/604,608
 ; PRIOR FILING DATE: 2000-06-27
 ; PRIOR APPLICATION NUMBER: 60/168,060
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: 60/177,836
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR APPLICATION NUMBER: 60/178,368
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: 60/210,292
 ; PRIOR FILING DATE: 2000-06-08
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Purified Memapsin 2
 ; OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide
 ; OTHER INFORMATION: residues
 ; OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,
 ; OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and
 ; OTHER INFORMATION: 376-377 are residues in contact with the OM99-2
 ; OTHER INFORMATION: inhibitor
 ; OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,
 ; OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and
 ; OTHER INFORMATION: 220-224 are N-lobe Beta Strands
 ; OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices
 ; OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,
 ; OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,
 ; OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,
 ; OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands
 ; OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,
 ; OTHER INFORMATION: and 427-431 are C-lobe Helices
 US-09-796-264-2

Query Match 96.9%; Score 2582; DB 10; Length 488;
 Best Local Similarity 99.8%; Pred. No. 2.4e-233;

Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 AGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEPEEPGRGSGFVEMVDNLRGKSGQ 73

DB 1 AGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEPEEPGRGSGFVEMVDNLRGKSGQ 60

QY 74 GYVEMTVGSPQTLNLVDGTSSNFVAGAAPFLHRYQRLSSTYRDLRGVVPYT 133

DB 61 GYVEMTVGSPQTLNLVDGTSSNFVAGAAPFLHRYQRLSSTYRDLRGVVPYT 120

QY 134 QKWEGELGTLVSIIPHGPNNVTVRANIAAITSDEKFFINGSNWEGILGLAYAEIARPDOS 193

DB 121 QKWEGELGTLVSIIPHGPNNVTVRANIAAITSDEKFFINGSNWEGILGLAYAEIARPDOS 180

QY 194 LEFFDLSLVKQTHVNPFLSHLCGAGFPLNQSEVLASVGSMTIGGDHSLYTGSLWYTP 253

DB 181 LEFFDLSLVKQTHVNPFLSHLCGAGFPLNQSEVLASVGSMTIGGDHSLYTGSLWYTP 240

QY 254 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSGTTNLRPKKVFEEAAVKSIAAS 313

DB 241 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSGTTNLRPKKVFEEAAVKSIAAS 300

QY 314 STEKFPDGFGLGEQLVCHQAGTTPWNIIPFVLSYLMGEVTVNQSFRTILPQOYLRPVEDV 373

DB 301 STEKFPDGFGLGEQLVCHQAGTTPWNIIPFVLSYLMGEVTVNQSFRTILPQOYLRPVEDV 360

QY 374 ATSDDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIKGFVAVSACHVHDEFTAAVEG 433

DB 361 ATSDDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIKGFVAVSACHVHDEFTAAVEG 420

QY 434 PFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVQWRCRLRQHQHDDF 493

DB 421 PFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVQWRCRLRQHQHDDF 480

QY 494 ADDISLLK 501

DB 481 ADDISLLK 488

RESULT 14

US-09-845-226-2

Sequence 2, Application US/09845226

Patent No. US20020115600A1

GENERAL INFORMATION:

APPLICANT: Tang, Jordan J.N.

APPLICANT: Hong, Lin

APPLICANT: Ghosh, Arun K.

TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof

FILE REFERENCE: OMRF 182

CURRENT APPLICATION NUMBER: US/09/845,226

PRIOR FILING DATE: 2001-04-30

PRIOR FILING DATE: 09/603,713

PRIOR FILING DATE: 2000-06-27

PRIOR FILING DATE: 60/168,060

PRIOR FILING DATE: 1999-11-30

PRIOR FILING DATE: 60/177,836

PRIOR FILING DATE: 2000-01-25

PRIOR FILING DATE: 60/178,368

PRIOR FILING DATE: 2000-01-27

PRIOR FILING DATE: 60/210,292

PRIOR FILING DATE: 2000-06-08

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 488

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Purified Memapsin 2

OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide

OTHER INFORMATION: residues

OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000

OTHER INFORMATION: 376-377 are residues in contact with the OM99-2

OTHER INFORMATION: inhibitor

OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111, 113-118, 123-134, 143-154, 165-168, 198-202, and 220-224 are N-lobe Beta Strands

OTHER INFORMATION: Amino acids 184-191 and 210-217 are N-lobe Helices

OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260, 273-275, 282-285, 316-318, 331-336, 342-348, 354-357, 366-370, 372-375, 380-383, 390-395, 400-405, and 418-420 are C-lobe Beta Strands

OTHER INFORMATION: Amino acids 286-299, 307-310, 350-353, 384-387, and 427-431 are C-lobe Helices

US-09-845-226-2

Query Match 96.9%; Score 2582; DB 10; Length 488;

Best Local Similarity 99.8%; Pred. No. 2.4e-233;

Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 AGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEPEEPGRGSGFVEMVDNLRGKSGQ 73

DB 1 AGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEPEEPGRGSGFVEMVDNLRGKSGQ 60

QY 74 GYVEMTVGSPQTLNLVDGTSSNFVAGAAPFLHRYQRLSSTYRDLRGVVPYT 133

DB 61 GYVEMTVGSPQTLNLVDGTSSNFVAGAAPFLHRYQRLSSTYRDLRGVVPYT 120

QY 134 QKWEGELGTLVSIIPHGPNNVTVRANIAAITSDEKFFINGSNWEGILGLAYAEIARPDOS 193

DB 121 QKWEGELGTLVSIIPHGPNNVTVRANIAAITSDEKFFINGSNWEGILGLAYAEIARPDOS 180

QY 194 LEFFDLSLVKQTHVNPFLSHLCGAGFPLNQSEVLASVGSMTIGGDHSLYTGSLWYTP 253

DB 181 LEFFDLSLVKQTHVNPFLSHLCGAGFPLNQSEVLASVGSMTIGGDHSLYTGSLWYTP 240

QY 254 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSGTTNLRPKKVFEEAAVKSIAAS 313

DB 241 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSGTTNLRPKKVFEEAAVKSIAAS 300

QY 314 STEKFPDGFGLGEQLVCHQAGTTPWNIIPFVLSYLMGEVTVNQSFRTILPQOYLRPVEDV 373

DB 301 STEKFPDGFGLGEQLVCHQAGTTPWNIIPFVLSYLMGEVTVNQSFRTILPQOYLRPVEDV 360

QY 374 ATSDDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIKGFVAVSACHVHDEFTAAVEG 433

DB 361 ATSDDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIKGFVAVSACHVHDEFTAAVEG 420

QY 434 PFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVQWRCRLRQHQHDDF 493

DB 421 PFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFMLPLCLMVQWRCRLRQHQHDDF 480

QY 494 ADDISLLK 501

DB 481 ADDISLLK 488

RESULT 15

US-09-795-903A-3

Sequence 3, Application US/09795903A

Patent No. US20020164760A1

GENERAL INFORMATION:

APPLICANT: Tang, Jordan J.N.

APPLICANT: Lin, Xinli

APPLICANT: Koelsch, Gerald

TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods of Use Thereof

FILE REFERENCE: OMRF 179

CURRENT APPLICATION NUMBER: US/09/795,903A

CURRENT FILING DATE: 2001-02-28

PRIOR APPLICATION NUMBER: 09/604,608

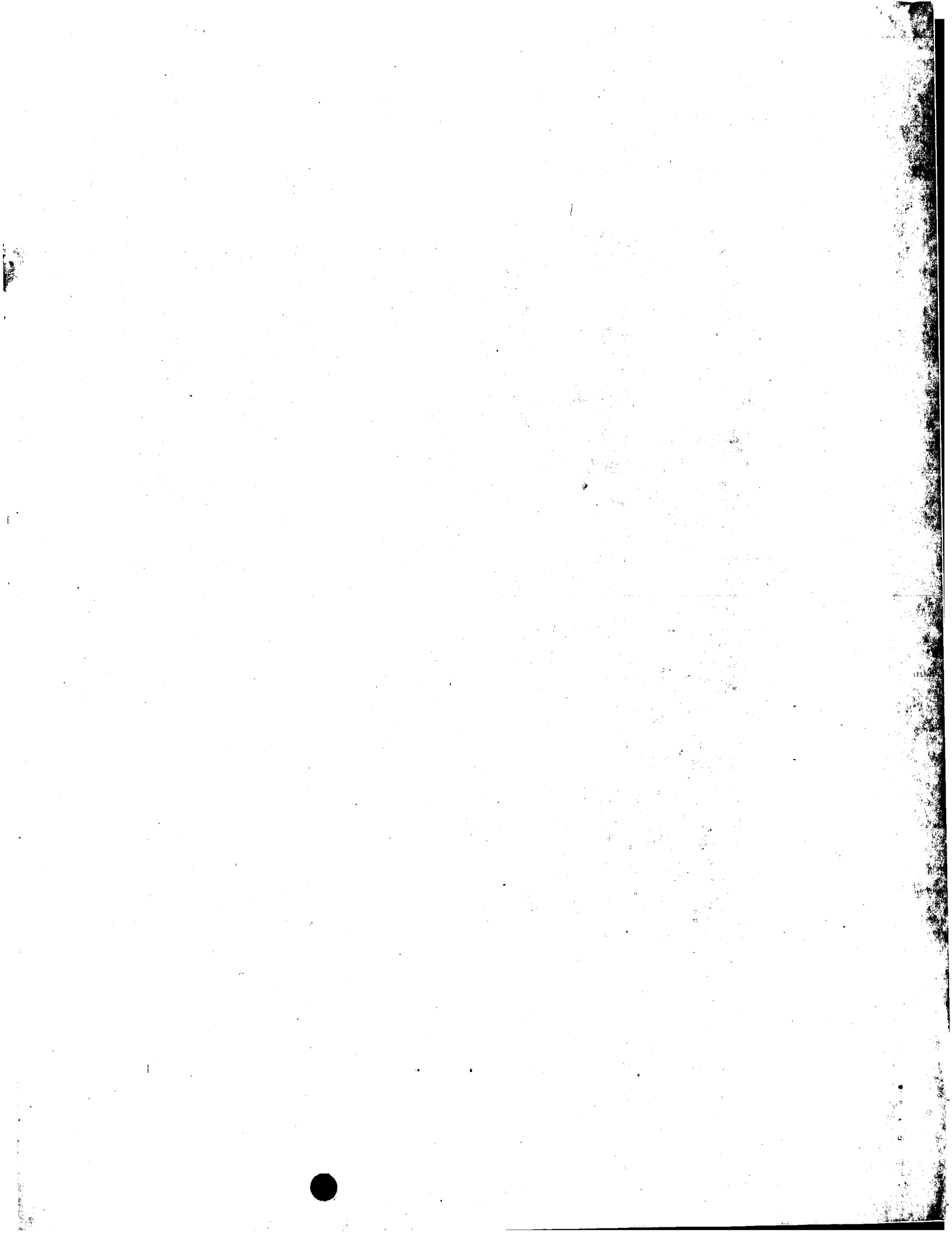
PRIOR FILING DATE: 2000-06-27

PRIOR APPLICATION NUMBER: 60/168,060

PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: 60/177,836

; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Pro-memapsin 2
; OTHER INFORMATION: Amino Acids 1-15 are vector-derived residues
; OTHER INFORMATION: Amino Acids 16-64 are a putative pro peptide
; OTHER INFORMATION: Amino Acids 1-13 are the T7 promoter
; OTHER INFORMATION: Amino Acids 16-456 are Pro-memapsin 2-T1
; OTHER INFORMATION: Amino Acids 16-421 are Promemapsin 2-T2
US-09-795-903A-3
Query Match 96.9%; Score 2582; DB 9; Length 503;
; Local Similarity 99.8%; Pred. No. 2.5e-233;
; Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 14 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPGRGSGFVEMVDNLRGKSGQ 73
Db 16 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPGRGSGFVEMVDNLRGKSGQ 75
Qy 74 GYVEMTVGSPPTQTLNLTVDGSSNFAVGAAPHPFLHRYRQRLSSYRDLRGVVPYT 133
Db 76 GYVEMTVGSPPTQTLNLTVDGSSNFAVGAAPHPFLHRYRQRLSSYRDLRGVVPYT 135
Qy 134 QGKWEGLGDLVSIHPGPNVTVRANIAATESDKFFINGSNWEGILGLAYAEIARPDSS 193
Db 136 QGKWEGLGDLVSIHPGPNVTVRANIAATESDKFFINGSNWEGILGLAYAEIARPDSS 195
Qy 194 LEFFFDLVKQTHVNPVLFSLHLCAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLWYTP 253
Db 196 LEFFFDLVKQTHVNPVLFSLHLCAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLWYTP 255
Qy 254 IRREWYEVIIIVVEINGQDLKWDCKEYNYDKSIVDSGTTNLRPLPKVFEAAVKSIAAS 313
Db 256 IRREWYEVIIIVVEINGQDLKWDCKEYNYDKSIVDSGTTNLRPLPKVFEAAVKSIAAS 315
Qy 314 STEKFPDGFGLGQLVCWQAGTTPWNIFFVISLYLMGEVNTNQSFRTILPQQYLRPVEDV 373
Db 316 STEKFPDGFGLGQLVCWQAGTTPWNIFFVISLYLMGEVNTNQSFRTILPQQYLRPVEDV 375
Qy 374 ATSQDDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFTAAVEG 433
Db 376 ATSQDDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFTAAVEG 435
Qy 434 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQWRCRLRQHQHDDF 493
Db 436 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQWRCRLRQHQHDDF 495
Qy 494 ADDISLLK 501
Db 496 ADDISLLK 503



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:03:10 ; Search time 38 Seconds
(without alignments)
1756.804 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLMGAGVLPAA.....CLRLRQHQHDFADISLLK 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2664	100.0	501	21	Human aspartyl pro
2	2664	100.0	501	22	Human aspartyl pro
3	2664	100.0	501	22	Human aspartyl pro
4	2664	100.0	501	22	Human aspartyl pro
5	2664	100.0	501	22	Human aspartyl pro
6	2664	100.0	501	22	Human aspartyl pro
7	2664	100.0	501	22	Human aspartyl pro
8	2664	100.0	501	23	Human aspartyl pro
9	2664	99.7	501	21	Human beta-secreta
10	2656	99.7	501	21	Amino acid sequenc

11	2656	99.7	509	23	AA52697	FLAG-tagged human
12	2650	99.5	501	19	AA59807	Amino acid sequenc
13	2588.5	97.2	969	22	ABG09611	Novel human diago
14	2582	96.9	488	22	AA666572	Human memapsin 2
15	2582	96.9	488	22	AA666572	Human memapsin 2
16	2582	96.9	488	23	AAU99488	Memapsin 2 protein
17	2582	96.9	501	22	AA84948	Human memapsin 2
18	2582	96.9	501	22	AA666573	Mouse aspartic sec
19	2582	96.9	503	22	AA666573	Human pro-memapsin
20	2582	96.9	503	22	AA666573	Human pro-memapsin
21	2569	96.4	501	21	AA94769	t7 promoter and ve
22	2567	96.4	501	21	AA94768	Pro-memapsin 2 enc
23	2567	96.4	501	21	AA94768	Rat beta-secretase
24	2567	96.4	501	21	AA94768	Murine beta-secret
25	2567	96.4	501	22	AAE10631	Murine aspartyl pr
26	2567	96.4	501	22	AAE10631	Murine aspartyl pr
27	2567	96.4	501	22	AAU06605	Murine aspartyl pr
28	2567	96.4	501	22	AAU07204	Mouse Aspartyl pro
29	2567	96.4	501	22	AAE02583	Mouse aspartyl pro
30	2506.5	94.1	476	21	AA84826	Murine aspartyl pr
31	2506.5	94.1	476	21	AA84826	Human aspartyl pr
32	2506.5	94.1	476	22	AAE10630	Human aspartyl pr
33	2506.5	94.1	476	22	AAE10630	Human aspartyl pr
34	2506.5	94.1	476	22	AAU07203	Human Aspartyl pro
35	2506.5	94.1	476	22	AAE02582	Human aspartyl pro
36	2506.5	94.1	476	23	ABB78591	Human Asp-2(b) pro
37	2506.5	94.1	476	23	ABB78591	Human aspartyl pr
38	2506.5	94.1	476	23	ABB06120	Human NS protein s
39	2420.5	90.9	476	22	AAE06909	Murine aspartyl pr
40	2420.5	90.9	476	22	AAE06909	Murine aspartyl pr
41	2414	90.6	456	21	AAE07897	Active enzyme port
42	2397	90.0	453	21	AA84838	Modified human asp
43	2397	90.0	453	21	AA84838	Human-Asp 2(a) pro
44	2397	90.0	453	22	AAE06872	Human-Asp2(a) delt
45	2397	90.0	453	22	AAU06616	Human-pro-Asp 2(a)

ALIGNMENTS

RESULT 1
AA588425
ID AA588425 standard; Protein; 501 AA.
XX
AC AA588425;
XX
DT 03-AUG-2000 (first entry)
XX
DE Human aspartyl protease 2 (a) (Asp2) amino acid sequence.
XX
KW Aspartyl protease; aspartase; amyloid precursor protein; APP; Asp 2;
KW Alzheimer's disease; beta secretase site.
XX
OS Homo sapiens.
XX
PN WO200017369-A2.
XX
PD 30-MAR-2000.
XX
PF 23-SEP-1999; 99WO-US20881.
XX
PR 24-SEP-1998; 98US-0101594.
XX
(PHAA) PHARMACIA & UPJOHN CO.
XX
PI Gurney ME, Bienkowski MJ, Heinrichson RL, Parodi LA, Yan R;
XX
DR WPI; 2000-303209/26.
XX
N-PSDB; AAA15662.
XX
PT New enzyme designated human aspartase useful in research into
PT Alzheimer's Disease is capable of cleaving amyloid protein precursor at
PT the beta secretase site to produce amyloid beta peptide

Human: aspartyl protease 2(a); Asp2(a); amyloid precursor protein; APP; Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis; amyloid plaque; neuronal loss; proteolytic; neurotrophic; neuroprotective; chromosome 11q23.3-24.1.

Homo sapiens.
Key Location/Qualifiers
1..21 /label= Signal_peptide
22..45 /label= Asp_2a_prepropeptide
46..57 /label= Asp_2a_propeptide
58..501 /label= Mature_human_Asp_2a_protein
420..454 /label= Alpha-helical_spacer_region
455..477 /label= Transmembrane_domain
478..501 /label= Cytoplasmic_domain

GB2357767-A.
04-JUL-2001.
22-SEP-2000; 2000GB-0023315.
23-SEP-1999; 99US-0155493.
23-SEP-1999; 99US-0404133.
23-SEP-1999; 99NO-US20881.
13-OCT-1999; 99US-0416901.
06-DEC-1999; 99US-0169232.
(PHAA) PHARMACIA & UPJOHN CO.
Bienkowski MJ, Gurney M;
WPI: 2001-444208/48.
N-PSDB: AAD17865.
Polypeptide comprising fragments of human aspartyl protease with amyloid precursor protein processing activity and alpha-secretase activity, for identifying modulators useful in treating Alzheimer's disease -
Example 2; Fig 2; 187pp; English.

The patent discloses human aspartyl protease 1 (hu-Asp1) or modified Asp1 proteins which lack transmembrane domain or amino terminal domain or cytoplasmic domain and retains alpha-secretase activity and amyloid protein precursor (APP) processing activity. The proteins of the invention are useful for assaying hu-Asp1 alpha-secretase activity, which in turn is useful for identifying modulators of hu-Asp1 alpha-secretase activity, where modulators that increase hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's disease (AD) which causes progressive dementia with consequent formation of amyloid plaques, neurofibrillary tangles, gliosis and neuronal loss. Hu-Asp1 protease substrate is useful for assaying hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with the substrate under acidic conditions and determining the level of hu-Asp1 proteolytic activity. The present sequence is long form of human Asp2 protein, designated as Asp2(a). Asp2 gene is localised on chromosome 11q23.3-24.1.

Query Match 100.0%; Score 2664; DB 22; Length 501;
Best Local Similarity 100.0%; Pred. No. 5,5e-263;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MAQALPWLILWMGAGVLPAGHTQHTGIRLPLRSLGGLGAPLGLPLRETDEEPEGRGSGF 60

XX PS Claim 48; Fig 2; 187pp; English.
XX This sequence represents the human aspartyl protease 2 (Asp2) amino acid sequence. The invention relates to a protease (e.g. Asp2) capable of cleaving the beta secretase site of amyloid precursor protein (APP). The protease contains a sequence encoding the amino acid sequence DRG and a sequence encoding DSG or DTG separated by 100-300 amino acids. When mutated the APP gene causes an autosomal dominant form of Alzheimer's disease. APP localises to the cell surface membrane and have a single C-terminal transmembrane domain. Proteolytic processing of APP produces the amyloid beta protein, which is possibly very important in Alzheimer's disease. The invention includes a nucleotide sequence, and a cell line comprising the vector. Methods for screening for inhibitors of beta secretase activity are also given in the invention. The human aspartase protein and nucleotide sequences and the methods for identifying inhibitors of the protease, are useful in the treatment of and research in to Alzheimer's disease.

Sequence 501 AA:
Query Match 100.0%; Score 2664; DB 21; Length 501;
Best Local Similarity 100.0%; Pred. No. 5,5e-263;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQALPWLILWMGAGVLPAGHTQHTGIRLPLRSLGGLGAPLGLPLRETDEEPEGRGSGF 60
Db 1 MAQALPWLILWMGAGVLPAGHTQHTGIRLPLRSLGGLGAPLGLPLRETDEEPEGRGSGF 60
QY 61 VEMVDNLKSGGGYVETVCSPPQTNLILVDGSSNFAYGCAAPHPHLHYRQRLSST 120
Db 61 VEMVDNLKSGGGYVETVCSPPQTNLILVDGSSNFAYGCAAPHPHLHYRQRLSST 120
QY 121 YRDLRGVYVPTQGWEGELGTLVSIPIHGPNTVTVRANIAITSDKFFINGSNWEGIL 180
Db 121 YRDLRGVYVPTQGWEGELGTLVSIPIHGPNTVTVRANIAITSDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLEFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDLEFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
QY 241 DHSLYTGLSWTPIRREYVEYIIVRVEINGQDLKMDCKEYNDKSYDVSQGTNLRLPKK 300
Db 241 DHSLYTGLSWTPIRREYVEYIIVRVEINGQDLKMDCKEYNDKSYDVSQGTNLRLPKK 300
QY 301 VFEAAVKSIIKAASSTKFPDGFVLGQVLCWQAGTTPWNIFPVISLYLMGEVTVNSFRIT 360
Db 301 VFEAAVKSIIKAASSTKFPDGFVLGQVLCWQAGTTPWNIFPVISLYLMGEVTVNSFRIT 360
QY 361 ILPOQYLRPVEDVATSDDCYKFAISQSSSTGTWNGAVIMEGFYVDFRARKRIGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDDCYKFAISQSSSTGTWNGAVIMEGFYVDFRARKRIGFAVSAC 420
QY 421 HVHDEFRTAAVEGFVFLDMEDCCYNTPORTDESTLMTIATVMAAICAFMLPCLMVCOW 480
Db 421 HVHDEFRTAAVEGFVFLDMEDCCYNTPORTDESTLMTIATVMAAICAFMLPCLMVCOW 480
QY 481 RCLRLCROQHDDFADDISLLK 501
Db 481 RCLRLCROQHDDFADDISLLK 501
RESULT 2
AAE10629
ID AAE10629 standard; Protein; 501 AA.
XX AAE10629;
AC AAE10629;
DT 10-DEC-2001 (first entry)
XX Human aspartyl protease 2(a) [hu-Asp2(a)] protein.
DE
XX

Db 1 MAQALPWLMMGAGVLPAGHTQHGIPLRLPSGLGAPLGLRLPRETDEPEEPPGRGSF 60
QY 61 VEMVDNLRKSGQGYVEMTVPQTLNILDVTGSSNFVAGAAPHPFLHRYYQRLSST 120
Db 61 VEMVDNLRKSGQGYVEMTVPQTLNILDVTGSSNFVAGAAPHPFLHRYYQRLSST 120
QY 121 YRDLRKGVPYVPTQKGWEGELGTDLVSPHGNVTVRANIAAITSDFKFFINGSNWEGIL 180
Db 121 YRDLRKGVPYVPTQKGWEGELGTDLVSPHGNVTVRANIAAITSDFKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240
QY 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVSQGTTLNLRPKK 300
Db 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVSQGTTLNLRPKK 300
QY 301 VFEAAVKSIKAASSTKFPDGLVGEOLVCWOAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
Db 301 VFEAAVKSIKAASSTKFPDGLVGEOLVCWOAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
QY 361 ILPOQYLRPEVDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKEIGFAVSAC 420
Db 361 ILPOQYLRPEVDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKEIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFPLCLMVQW 480
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFPLCLMVQW 480
QY 481 RCLRLRQOHHDFADDISLLK 501
Db 481 RCLRLRQOHHDFADDISLLK 501

RESULT 3

AAE06859
ID AAE06859 standard; Protein; 501 AA.

XX AC AAE06859;

XX DT 23-OCT-2001 (first entry)

XX DE Human aspartyl protease 2a (Hu-Asp2a) protein.

XX KW Human; aspartyl protease 2a; Asp 2a; beta-amyloid precursor protein; App;
KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nontropic;
KW neuroprotective; antisense therapy; gene therapy;
XX chromosome 11q23.3-24.1.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..21

FT Protein /label= signal_peptide

FT Region /note= "Mature human aspartyl protease 2a (Hu-Asp2a)"

FT Domain /note= "Alpha helical spacer region"

FT Domain /label= Transmembrane_domain

FT Domain /label= Cytoplasmic_domain

XX WO200150829-A2.

XX 19-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00799.

XX 09-MAY-2001; 2001WO-IB00799.

XX PA (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX DR WPI: 2001-483072/52.
XX DR N-PSDB; AAD13021.
XX PT Novel purified polypeptide comprising fragment of mammalian aspartyl
XX PT protease 2, lacking Asp2 transmembrane domain and retaining beta
XX PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
XX PS activity
XX PS Claim 49; Fig 2; 185pp; English.
XX CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
XX CC precursor protein (APP) isoforms and their corresponding DNA molecules.
XX CC Human aspartyl proteases can act as beta-secretase proteases useful for
XX CC treating Alzheimer's disease. APP isoforms are useful for identifying
XX CC modulators of amyloid-beta peptide production, for use in designing
XX CC therapeutics for the treatment and prevention of Alzheimer's disease,
XX CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
XX CC and neuronal loss. APP isoforms are also used in methods for identifying
XX CC inhibitors and modulators of human Asp2 activity. The invention relates
XX CC to a method for identifying agents that modulate the activity of human
XX CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
XX CC as a means to screen in cellular assays for the inhibitors of beta- and
XX CC gamma- secretase: Hu-Asp DNA fragments are useful as probes or primers in
XX CC polymerase chain reactions (PCR). The probes are useful for detecting
XX CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
XX CC blots. The present sequence is human aspartyl protease 2 (Hu-Asp2), a
XX CC 'long' form designated as (Hu-Asp2a). Hu-Asp 2 gene is localised on
XX CC chromosome 11q23.3-24.1.
XX SQ Sequence 501 AA;

Query Match 100.0%; Score 2664; DB 22: Length 501;

Best Local Similarity 100.0%; Pred. No. 5.5e-263;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQALPWLMMGAGVLPAGHTQHGIPLRLPSGLGAPLGLRLPRETDEPEEPPGRGSF 60
Db 1 MAQALPWLMMGAGVLPAGHTQHGIPLRLPSGLGAPLGLRLPRETDEPEEPPGRGSF 60
QY 61 VEMVDNLRKSGQGYVEMTVPQTLNILDVTGSSNFVAGAAPHPFLHRYYQRLSST 120
Db 61 VEMVDNLRKSGQGYVEMTVPQTLNILDVTGSSNFVAGAAPHPFLHRYYQRLSST 120
QY 121 YRDLRKGVPYVPTQKGWEGELGTDLVSPHGNVTVRANIAAITSDFKFFINGSNWEGIL 180
Db 121 YRDLRKGVPYVPTQKGWEGELGTDLVSPHGNVTVRANIAAITSDFKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPLNLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240
QY 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVSQGTTLNLRPKK 300
Db 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVSQGTTLNLRPKK 300
QY 301 VFEAAVKSIKAASSTKFPDGLVGEOLVCWOAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
Db 301 VFEAAVKSIKAASSTKFPDGLVGEOLVCWOAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
QY 361 ILPOQYLRPEVDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKEIGFAVSAC 420
Db 361 ILPOQYLRPEVDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKEIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALFPLCLMVQW 480

APP or its fragment containing an APP cleavage site recognizable by a mammalian beta-secretase, and further comprising two lysine residues at the carboxyl terminus of the amino acid sequence of the mammalian APP or APP fragment. Also included in the invention are methods of identifying modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are useful for treating Alzheimer's disease. APP is useful in methods for identifying inhibitors or modulators of human Asp2 activity and amyloid-beta (Abeta) peptide production. APP is also useful in designing therapeutics for the treatment or prevention of Alzheimer's disease. APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which is associated with increased levels of Abeta processing is useful in assays relating the Alzheimer's research. The expression vector is useful for recombinantly expressing APP. Nucleic acids that hybridize to Asp oligonucleotides are useful as probes or primers. The probes are useful for detecting Hu-Asp nucleic acids in in vitro assays and in Northern and Southern blots. The present sequence is human Asp2(a).

CC APP or its fragment containing an APP cleavage site recognizable by a
CC mammalian beta-secretase, and further comprising two lysine residues at
CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
CC APP fragment. Also included in the invention are methods of identifying
CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
CC useful for treating Alzheimer's disease. APP is useful in methods for
CC identifying inhibitors or modulators of human Asp2 activity and
CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
CC therapeutics for the treatment or prevention of Alzheimer's disease.
CC APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which
CC is associated with increased levels of Abeta processing is useful in
CC assays relating the Alzheimer's research. The expression vector is useful
CC for recombinantly expressing APP. Nucleic acids that hybridize to
CC Asp oligonucleotides are useful as probes or primers. The probes are
CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC Northern and Southern blots. The present sequence is human Asp2(a).
XX SQ Sequence 501 AA;

Query Match 100.0%; Score 2664; DB 22; Length 501;
Best Local Similarity 100.0%; Pred. No. 5,5e-263;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMGAGVLPAGHCTGIRLPLRSGLGGAPLGLRLPRETDEPEEPGRGSF 60
DB 1 MAQALPWLMLMGAGVLPAGHCTGIRLPLRSGLGGAPLGLRLPRETDEPEEPGRGSF 60
QY 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNILVDVTGSSNFVAGAPHPFLHRYQRLSST 120
DB 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNILVDVTGSSNFVAGAPHPFLHRYQRLSST 120
QY 121 YDLRKGYYVPTQGWEGELGTDLVSPHGPNTVVRANTAAITSDKFFINGSNWEGIL 180
DB 121 YDLRKGYYVPTQGWEGELGTDLVSPHGPNTVVRANTAAITSDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVFNLSHLCGAGPPLNQSEVLASVGGSMIIGGI 240
DB 181 GLAYAEIARPDSDLPPFDLSLVKQTHVFNLSHLCGAGPPLNQSEVLASVGGSMIIGGI 240
QY 241 DHSLYTGSLLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIIVDSCTTNLRLPKK 300
DB 241 DHSLYTGSLLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIIVDSCTTNLRLPKK 300
QY 301 VFEAAVKSIIKAASSTKPKDPGFWLGEOLVCWQAGTTPWNIIPVISLYLMGEVTVNQSPRIT 360
DB 301 VFEAAVKSIIKAASSTKPKDPGFWLGEOLVCWQAGTTPWNIIPVISLYLMGEVTVNQSPRIT 360
QY 361 ILPQOYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
DB 361 ILPQOYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
QY 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFPLCLMVQW 480
DB 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFPLCLMVQW 480
QY 481 RCLRLRQQHDDFADDSILK 501
DB 481 RCLRLRQQHDDFADDSILK 501

RESULT 5
AAU07202
ID AAU07202 standard; Protein: 501 AA.
XX
XX AAU07202;
XX 24-OCT-2001 (first entry)
XX Human aspartyl protease 2a (Asp-2a).
XX Human aspartyl protease 1; Asp-1; nootropic; neuroprotective;
XX aspartyl protease 2; Asp2; amyloid protein precursor; APP;
XX beta-secretase; Alzheimer's disease.

DB 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFPLCLMVQW 480
QY 481 RCLRLRQQHDDFADDSILK 501
DB 481 RCLRLRQQHDDFADDSILK 501

RESULT 4
AAU06603
ID AAU06603 standard; Protein: 501 AA.

XX

XX

XX 24-OCT-2001 (first entry)

XX Human Aspartyl protease 2(a), Asp2(a).

XX Human; Aspartyl protease; Asp2(a); beta-secretase; nootropic;
XX neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
XX amyloid-beta; Abeta.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..21 /label= Signal peptide

XX Peptide 22..45 /label= Pre_pro-peptide

XX Peptide 46..57 /label= Pro-peptide

XX Protein 57..501 /label= Mature_Asp2(a)

XX Region 420..454 /label= Alpha_helical_region

XX Domain 455..477 /label= Transmembrane_domain

XX Domain 478..501 /label= Cytoplasmic_domain

XX WO200149098-A2.

XX 12-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00798.

XX 09-MAY-2001; 2001WO-IB00798.

XX (BIEN/) BIENKOWSKI M J.

XX (GURN/) GURNEY M E.

XX (HEIN/) HEINRIKSON R L.

XX (PARO/) PARODI L A.

XX (YANR/) YAN R.

XX Bienkowski MJ, Gurney ME, Henrikson RL, Parodi LA, Yan R;

XX WPI: 2001-502549/55.

XX N-PSDB; AAS11517.

XX Novel purified polypeptide comprising fragment of mammalian aspartyl

XX protease 2, lacking Asp2 transmembrane domain and retaining beta

XX secretase activity of Asp2 useful for identifying inhibitors of Asp2

XX activity

XX Claim 49; Fig 2; 185pp; English.

XX The invention relates to a purified polypeptide comprising a fragment of

XX mammalian aspartyl protease (Asp2) protein which lacks the Asp2

XX transmembrane domain and the Asp2 protein, and where the polypeptide and

XX the fragment retain the beta-secretase activity of the mammalian Asp2

XX protein. The invention also details polynucleotides for the Asp

XX proteins and vectors expressing them, and a polypeptide (isoform of

XX amyloid protein precursor (APP)) comprising the amino acid sequence of an

XX	OS	Homo sapiens.	
XX	FT	Key	Location/Qualifiers
XX	FT	Peptide	1..21
XX	FT	Misc_feature	/note= "Signal peptide"
XX	FT	Misc_feature	22..45
XX	FT	Misc_feature	/note= "Pre-propeptide"
XX	FT	Misc_feature	46..57
XX	FT	Protein	/note= "Propeptide"
XX	FT	Protein	58..501
XX	FT	Region	/note= "Mature Aspartyl protease-2a"
XX	FT	Region	420..454
XX	FT	Domain	/note= "Alpha helical spacer region"
XX	FT	Domain	455..477
XX	FT	Domain	/note= "Transmembrane domain"
XX	FT	Domain	478..501
XX	FT	Domain	/note= "Cytoplasmic domain"
XX	PN	WO200149097-A2.	
XX	XX	12-JUL-2001.	
XX	PF	09-MAY-2001; 2001WO-IB00797.	
XX	PR	09-MAY-2001; 2001WO-IB00797.	
XX	PA	(BIEN/) BIENKOWSKI M J.	
XX	PA	(GURN/) GURNEY M E.	
XX	PA	(HEIN/) HEINRIKSON R L.	
XX	PA	(PARO/) PARODI L A.	
XX	PA	(YANR/) YAN R.	
XX	PI	Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;	
XX	DR	WPI: 2001-502538/55.	
XX	DR	N-PSDB: AAS11702.	
XX	PT	Novel purified polypeptide comprising fragment of mammalian aspartyl	
XX	PT	protease 2, lacking Asp2 transmembrane domain and retaining beta	
XX	PT	secretase activity of Asp2 useful for identifying inhibitors of Asp2	
XX	PT	activity -	
XX	PS	Claim 49; Fig 2; 185pp; English.	
XX	CC	The invention relates to a novel purified polypeptide comprising a	
XX	CC	fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the	
XX	CC	Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide	
XX	CC	and the fragment retain the beta-secretase activity of the mammalian Asp2	
XX	CC	protein. Also included is an isoform of amyloid precursor protein (APP)	
XX	CC	comprising the amino acid sequence of a APP or its fragment containing	
XX	CC	an APP cleavage site recognisable by a mammalian beta-secretase, and	
XX	CC	further comprising two lysine residues at the carboxyl terminus of the	
XX	CC	amino acid sequence of the mammalian APP or APP fragment. The	
XX	CC	polypeptides are used for assaying for modulators of beta-secretase	
XX	CC	activity; identifying agents that inhibit the APP processing activity	
XX	CC	of human Asp2 aspartyl protease (Hu-Asp2); identifying agents that	
XX	CC	modulate the activity of Asp2; and for reducing cellular production of	
XX	CC	amyloid beta (Abeta) from APP. Agents identified by the above methods	
XX	CC	are useful for treating Alzheimer's disease; and for identifying	
XX	CC	modulators of amyloid-beta (Abeta) peptide production, for use in	
XX	CC	designing therapeutics for the treatment or prevention of Alzheimer's	
XX	CC	disease. Probes and primers derived from Asp nucleic acid sequences	
XX	CC	are useful for detecting Hu-Asp nucleic acids in vitro assays and in	
XX	CC	Northern and Southern blots. The present sequence represents the	
XX	CC	amino acid sequence of human Asp-2a used in the methods of the invention.	
XX	SQ	Sequence 501 AA;	
XX	XX	Query Match	100.0%; Score 2664; DB 22; Length 501;
XX	XX	Best Local Similarity	100.0%; Pred. No. 5.5e-263;
XX	XX	Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

Qy	1	MAQALPWLMMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF	60
Db	1	MAQALPWLMMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF	60
Qy	61	VEVDNLGRKSGQCYVVEVTGSPQTLNLLVDGSSNFAVGAAPHPFLHRYYQRLSST	120
Db	61	VEVDNLGRKSGQCYVVEVTGSPQTLNLLVDGSSNFAVGAAPHPFLHRYYQRLSST	120
Qy	121	YRDLRKGVYVPTQGWEGELGTDLVSTPHGPNVTVRANIAAITESDKFFINGSNWEGIL	180
Db	121	YRDLRKGVYVPTQGWEGELGTDLVSTPHGPNVTVRANIAAITESDKFFINGSNWEGIL	180
Qy	181	GLAYAEIARPPDLSLEPFDSLVKQTHVFNLSLHLCGAGFPPLNQSSEVLASVGGSMIIGGI	240
Db	181	GLAYAEIARPPDLSLEPFDSLVKQTHVFNLSLHLCGAGFPPLNQSSEVLASVGGSMIIGGI	240
Qy	241	DHSLYTGLWYTPIRREWIYEVIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK	300
Db	241	DHSLYTGLWYTPIRREWIYEVIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK	300
Qy	301	VFEAAVKSIIKAASSTKFPDGLGEQVLCVQAGTTPWNIFPVISLSYLMGEVTTNQSFRIT	360
Db	301	VFEAAVKSIIKAASSTKFPDGLGEQVLCVQAGTTPWNIFPVISLSYLMGEVTTNQSFRIT	360
Qy	361	ILPOQYLRPVEDVATSDQDCYKFAISQSSSTGTVMGAVIMEGFYVYVDFRARRIGFAVSAC	420
Db	361	ILPOQYLRPVEDVATSDQDCYKFAISQSSSTGTVMGAVIMEGFYVYVDFRARRIGFAVSAC	420
Qy	421	HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDESTLMTIAYVMAATCALFPLCLMVCOW	480
Db	421	HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDESTLMTIAYVMAATCALFPLCLMVCOW	480
Qy	481	RCLRLRQHQHDDFADDISLLK 501	
Db	481	RCLRLRQHQHDDFADDISLLK 501	
XX	RESULT 6		
XX	AAE02581	standard; Protein; 501 AA.	
XX	AC	AAE02581;	
XX	DT	10-AUG-2001 (first entry)	
XX	DE	Human aspartyl protease 2a (Asp 2a).	
XX	KW	Human; alpha-secretase; amyloid precursor protein; APP; therapy;	
XX	KW	Alzheimer's disease; antialzheimer's; aspartyl protease 2a; Asp 2a;	
XX	OS	beta-secretase; chromosome 11q23.3-24.1.	
XX	OS	Homo sapiens.	
XX	FT	Key	Location/Qualifiers
XX	FT	Peptide	1..21
XX	FT	Peptide	/label= Signal_peptide
XX	FT	Peptide	22..45
XX	FT	Peptide	/label= Asp_2a_prepropeptide
XX	FT	Peptide	46..57
XX	FT	Peptide	/label= Asp_2a_propeptide
XX	FT	Protein	58..501
XX	FT	Protein	/label= Mature_human_Asp_2a_protein
XX	FT	Active-site	93..95
XX	FT	Active-site	/label= Active_site_1
XX	FT	Active-site	289..291
XX	FT	Active-site	/label= Active_site_2
XX	FT	Region	420..454
XX	FT	Region	/label= Alpha_helical_spacer
XX	FT	Domain	455..477
XX	FT	Domain	/label= Transmembrane_domain
XX	FT	Domain	478..501
XX	FT	Domain	/label= Cytoplasmic_domain
XX	FT	Region	486..501

FT WO200123533-A2. /note= "Peptide #2"

PN XX

PD 05-APR-2001.

XX 22-SEP-2000; 2000WO-US26080.

PF 23-SEP-1999; 99US-0155493.

PR 23-SEP-1999; 99WO-US20881.

PR 13-OCT-1999; 99US-0416901.

PR 06-DEC-1999; 99US-0169232.

XX (PHAA) PHARMACIA & UPJOHN CO.

PA Gurney M, Bienkowski MJ;

XX WPI; 2001-290516/30.

PI N-PSDB; AAD06739.

DR

PT Enzymes that cleave the alpha-secretase site of the amyloid precursor protein, useful for the treatment of Alzheimer's disease -

XX Example 2; Fig 2; 189pp; English.

PS The present invention relates to enzymes for cleaving the alpha-secretase site of the amyloid precursor protein (APP) and methods of identifying those enzymes. The methods may be used to identify enzymes that may be used to cleave the alpha-secretase cleavage site of the APP protein. The enzymes may be used to treat or modulate the progress of Alzheimer's disease. The present sequence is human aspartyl protease 2a (Asp 2a). Asp 2a has beta-secretase protease activity. Asp 2 gene is located on chromosome 11q23.3-24.1.

XX Sequence 501² AA;

SQ Query Match 100.0%; Score 2664; DB 22; Length 501;

Best Local Similarity 100.0%; Pred. No. 5.5e-263;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMGAGVLPAGHTGIRLPLRSLGAGPLGLRLPRETDEEPEERGRGSF 60

DB 1 MAQALPWLMLMGAGVLPAGHTGIRLPLRSLGAGPLGLRLPRETDEEPEERGRGSF 60

QY 61 VEMVDNLRGSGGCGYVEMTVGSPPTNLILVDGSSNFVAGAAPHPFLHRYYQRLSST 120

DB 61 VEMVDNLRGSGGCGYVEMTVGSPPTNLILVDGSSNFVAGAAPHPFLHRYYQRLSST 120

QY 121 YRDLRKGVYVPYTGQKWEGLGDLVSIPIHGPNTVVRANTAAITSDKFFINGSNWEGIL 180

DB 121 YRDLRKGVYVPYTGQKWEGLGDLVSIPIHGPNTVVRANTAAITSDKFFINGSNWEGIL 180

QY 181 GLAYAEATRPDLSLEPFDSLVKQTHVPNLFSLHLCGAGPPLNQSEVLASVSGSMIIGGI 240

DB 181 GLAYAEATRPDLSLEPFDSLVKQTHVPNLFSLHLCGAGPPLNQSEVLASVSGSMIIGGI 240

QY 241 DLSLYTGSLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

DB 241 DLSLYTGSLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

QY 301 VFEAAVKSIKAASSTKPPDGFVWLGEOLVCHQAGTTPNIPVLSYLMGEVNTNQSFRT 360

DB 301 VFEAAVKSIKAASSTKPPDGFVWLGEOLVCHQAGTTPNIPVLSYLMGEVNTNQSFRT 360

QY 361 ILPQOYLARVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKRUFVAVSAC 420

DB 361 ILPQOYLARVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKRUFVAVSAC 420

QY 421 HVHDEFRTAAVEGPPVTLLDMEDCGYNIPQDDESTLMTIAYVMAAICALFMLPLCLMWQW 480

DB 421 HVHDEFRTAAVEGPPVTLLDMEDCGYNIPQDDESTLMTIAYVMAAICALFMLPLCLMWQW 480

QY 481 RCLRLRQOHHDDFADDISLLK 501

DB 481 RCLRLRQOHHDDFADDISLLK 501

RESULT 7

ABB78590

ID ABB78590 standard; Protein; 501 AA.

XX ABB78590;

AC ABB78590;

DT 16-JUL-2002 (first entry)

XX Human Asp-2(a) protein sequence SEQ ID NO:4.

XX Human; Asp-1; Asp-2; aspartyl protease; enzyme; Alzheimer's disease;

KW proteolytic; chromosome 11q23.3-24.1.

XX Homo sapiens.

OS GB2367060-A.

PN 27-MAR-2002.

PD 29-OCT-2001; 2001GB-0025934.

XX 23-SEP-1999; 99US-155493P.

PR 23-SEP-1999; 99US-0404133.

PR 23-SEP-1999; 99WO-US20881.

PR 13-OCT-1999; 99US-0416901.

PR 06-DEC-1999; 99US-169232P.

PR 22-SEP-2000; 2000GB-0023315.

XX (PHAA) PHARMACIA & UPJOHN CO.

XX Bienkowski MJ, Gurney M;

XX WPI; 2002-396337/43.

XX N-PSDB; ABL52457.

PT Human aspartyl protease 1 substrates useful in assays to detect

PT aspartyl protease activity, e.g. for the diagnosis of Alzheimer's

XX disease -

PS Example 2; Fig 2; 182pp; English.

XX The present invention describes a human aspartyl protease 1 (hu-Asp1)

XX substrate (I) which comprises a peptide of no more than 50 amino acids,

XX and which comprises the 8 amino acid sequence Gly-Leu-Ala-Leu-Ala-Leu-

XX Glu-Pro. Also described are: (1) a method (II) for assaying hu-Asp1

XX proteolytic activity, comprising: (a) contacting a hu-Asp1 protein with

XX (I) under acidic conditions; and (b) determining the level of hu-Asp1

XX proteolytic activity; (2) a purified polynucleotide (III) comprising a

XX nucleotide sequence that hybridises under stringent conditions to the

XX non-coding strand complementary to a defined 1804 nucleotide sequence

XX (see ABL52456) where the nucleotide sequence encodes a polypeptide having

XX Asp1 proteolytic activity and lacks nucleotides encoding a transmembrane

XX domain; (3) a purified polynucleotide (III') comprising a sequence that

XX hybridises under stringent conditions to (III) (the nucleotide sequence

XX encodes a polypeptide further lacking a pro-peptide domain corresponding

XX to amino acids 23-62 of hu-Asp1 (see ABB78589)); (4) a vector (IV)

XX comprising (III) or (III') and (5) a host cell (V) transformed or

XX transfected with (III), (III') and/or (IV). The hu-Asp1 protease

XX substrate (I) may be used as an enzyme substrate in assays to detect

XX aspartyl protease activity, (II) and therefore diagnose diseases

XX associated with aberrant hu-Asp1 expression and activity such as

XX Alzheimer's disease. Hu-Asp1 has been localised to chromosome 21, while

XX hu-Asp2 has been localised to chromosome 11q23.3-24.1. The present

XX sequence represents hu-Asp2(a) from the present invention.

XX Sequence 501 AA;

SQ Query Match 100.0%; Score 2664; DB 23; Length 501;

Best Local Similarity 100.0%; Pred. No. 5.5e-263;

Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease; human;
Down's syndrome; amyloid angiopathy; gene therapy; neuroprotective.

Homo sapiens.

Key Location/Qualifiers
Peptide 1..45
Protein /label= putative signal peptide
/label= Beta-secretase

WO200058479-A1.
05-OCT-2000.
23-MAR-2000; 2000WO-US07755.
26-MAR-1999; 99US-0277229.
(AMGE-) AMGEN INC.

Citron M, Vassar RJ, Bennett BD;
WPI; 2000-594643/56.
N-PSDB; AAA28278.

Isolated beta-secretase nucleic acids and encoded polypeptides, useful
for diagnosis and gene therapy of Alzheimer's disease

Claim 1; Fig 4; 145pp; English.

This invention relates to 3 nucleotide sequences encoding beta-secretase
proteins. Beta-secretase is an enzyme involved in the production of one
of the components of amyloid plaques involved in Alzheimer's disease. The
invention includes an expression vector comprising the nucleotide
sequence, a host cell comprising the expression vector, and a process for
producing the protein through culturing the transformed cells. Also
included in the invention are a polypeptide derivative of the
beta-secretase protein, a fusion protein comprising beta-secretase fused
to a heterologous amino acid sequence, and a method for modulating the
levels of beta-secretase polypeptide in a mammal comprising administering
the polynucleotide sequence. Beta-secretase exhibits neuroprotective and
neurotropic activity. The beta-secretase nucleotide sequence may be used to
map locations of the beta-secretase gene and related genes on chromosomes
and as hybridization probes in diagnostic assays to test for the presence
of beta-secretase DNA or RNA, such as in Alzheimer's disease, Down's
syndrome, and amyloid angiopathy. The nucleotide sequence may also be
used as anti-sense inhibitors of beta-secretase expression, in gene
therapy of Alzheimer's disease, and for the identification of compounds
that modulate beta-secretase activity. Antibodies to the beta-secretase
protein may be used for in vitro and in vivo diagnostic purposes to
detect the presence of beta-secretase polypeptide in a body fluid or cell
sample. The present sequence represents the human beta-secretase protein.

Sequence 501 AA:

Query Match 99.7%; Score 2656; DB 21; Length 501;
Best Local Similarity 99.8%; Pred No. 3.6e-262;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQALPWLKMGAGVLPARGTQHGIRPLRSLGAGAPLGLRLPRETDEEPEEPCRGGSF 60
DB 1 MAQALPWLKMGAGVLPARGTQHGIRPLRSLGAGAPLGLRLPRETDEEPEEPCRGGSF 60
QY 61 VEMVDNLKSGGGYVEMTVGSPPTNLIVDTGSSNFVAGAAPHPFLHRYQRQLSST*120
DB 61 VEMVDNLKSGGGYVEMTVGSPPTNLIVDTGSSNFVAGAAPHPFLHRYQRQLSST 120
QY 121 YRDLRKGVPYPTQGWEGELGTDLVSIIPHGNVTVVRANAAITSDKFFINGSNWEGIL 180
DB 121 YRDLRKGVPYPTQGWEGELGTDLVSIIPHGNVTVVRANAAITSDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLLEFFDLSLVKQTHVPLNLSLQCGAGFPLNQSEVLASVGGSMIIGGI 240

DB 181 GLAYAEIARPDSDLLEFFDLSLVKQTHVPLNLSLQCGAGFPLNQSEVLASVGGSMIIGGI 240
QY 241 DHSLYTGSLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIYVDSGTTNLRLPKK 300
DB 241 DHSLYTGSLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIYVDSGTTNLRLPKK 300
QY 301 VFEAAVKSITKAASSTKFPDGFGLGOLYCVQWAGTTPWNIFPVISLYLMGEVTVNQSFRIT 360
DB 301 VFEAAVKSITKAASSTKFPDGFGLGOLYCVQWAGTTPWNIFPVISLYLMGEVTVNQSFRIT 360
QY 361 ILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVYVDFDRARKRIGFAVSAC 420
DB 361 ILPQOYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVYVDFDRARKRIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDTESTLMTIAYVMAAICALFMLPLCLMVCQW 480
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDTESTLMTIAYVMAAICALFMLPLCLMVCQW 480
QY 481 RCLRLRQOHHDFADDISLLK 501
DB 481 RCLRLRQOHHDFADDISLLK 501

RESULT 10
AAB07896
ID AAB07896 standard; Protein; 501 AA.
XX
AC AAB07896;

DT 14-NOV-2000 (first entry)

XX Amino acid sequence of a human beta-secretase enzyme.

DE Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.

XX Homo sapiens.

XX WO2000047618-A2.

XX 17-AUG-2000.

XX 10-FEB-2000; 2000WO-US03819.

XX 10-FEB-1999; 99US-0119571.

XX 15-JUN-1999; 99US-0139172.

XX (ELAN-) ELAN PHARM INC.

XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;

XX Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;

XX WPI; 2000-533011/48.

XX N-PSDB; AAA59550, AAA59551.

XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease

XX Claim 17; Fig 2A; 121pp; English.

XX The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a human beta-secretase enzyme.

XX	DB	361	ILPQOYLRPVEDVATSDQCYKFAISQSSSTGTVMGAVIMEGFYVDFDRARRKIGFAVSAC	420
XX	QY	421	HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAAICALFMLPLCLMVCOW	480
XX	DB	421	HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAAICALFMLPLCLMVCOW	480
XX	QY	481	RCLRCLRQHQHDDFADDISLLK	501
XX	DB	481	RCLRCLRQHQHDDFADDISLLK	501
XX	RESULT 13			
XX	ID	ABG09611	standard; Protein: 969 AA.	
XX	AC	ABG09611;		
XX	DT	13-FEB-2002	(first entry)	
XX	XX	Novel human diagnostic protein #9602.		
XX	DE	Human: chromosome mapping; gene mapping; gene therapy; forensic;		
XX	KW	food supplement; medical imaging; diagnostic; genetic disorder.		
XX	KW	Homo sapiens.		
XX	OS	WG200175067-A2.		
XX	PN	11-OCT-2001.		
XX	PD	30-MAR-2001; 2001WO-US08631.		
XX	PF	31-MAR-2000; 2000US-0540217.		
XX	PR	23-AUG-2000; 2000US-0649167.		
XX	XX	(HYSE-) HYSEQ INC.		
XX	PA	Drmanac RT, Liu C, Tang YT;		
XX	PI	WPI; 2001-639362/73.		
XX	DR	N-PSDB; AAS73798.		
XX	DR	New isolated polynucleotide and encoded polypeptides, useful in		
XX	PT	diagnostics, forensics, gene mapping, identification of mutations		
XX	PT	responsible for genetic disorders or other traits and to assess		
XX	PT	biodiversity		
XX	XX	Claim 20; SEQ ID NO 39970; 103pp; English.		
XX	XX	The invention relates to isolated polynucleotide (I) and		
XX	CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,		
XX	CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome		
XX	CC	and gene mapping, and in recombinant production of (II). The		
XX	CC	polynucleotides are also used in diagnostics as expressed sequence tags		
XX	CC	for identifying expressed genes. (I) is useful in gene therapy techniques		
XX	CC	to restore normal activity of (II) or to treat disease states involving		
XX	CC	quantitating a polypeptide in tissue, as molecular weight markers and as		
XX	CC	a food supplement. (II) and its binding partners are useful in medical		
XX	CC	imaging of sites expressing (II). (I) and (II) are useful for treating		
XX	CC	disorders involving aberrant protein expression or biological activity.		
XX	CC	The polypeptide and polynucleotide sequences have applications in		
XX	CC	diagnostics, forensics, gene mapping, identification of mutations		
XX	CC	responsible for genetic disorders or other traits to assess biodiversity		
XX	CC	and to produce other types of data and products dependent on DNA and		
XX	CC	amino acid sequences. ABG00010-ABG30377 represent novel human		
XX	CC	diagnostic amino acid sequences of the invention.		
XX	CC	Note: The sequence data for this patent did not appear in the printed		
XX	CC	specification, but was obtained in electronic format directly from WIPO		
XX	CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX	XX	Sequence 969 AA;		
XX	SQ			
XX	DB	26-OCT-1998	(first entry)	
XX	XX	Amino acid sequence of human ASP2 (aspartic protease 2).		
XX	XX	Human: ASP2; aspartic protease 2; agonist; antagonist; immunospecific;		
XX	KW	antibody; inhibition; Alzheimer's disease; cancer; proteinase;		
XX	KW	prohormone processing.		
XX	OS	Homo sapiens.		
XX	EP	EP855444-A2.		
XX	PD	29-JUL-1998.		
XX	PF	27-JAN-1998; 98EP-0300573.		
XX	PR	28-JAN-1997; 97GB-0001684.		
XX	XX	(SMIK) SMITHKLINE BEECHAM CORP.		
XX	XX	(SMIK) SMITHKLINE BEECHAM PLC.		
XX	PI	Chapman CG, Murphy K, Powell DJ, Smith TS;		
XX	WPI	1998-389809/34.		
XX	N-PSDB	AAV41696.		
XX	XX	New nucleic acid encoding human aspartic protease 2 - used to treat,		
XX	PT	prevent and diagnose e.g. Alzheimer's disease, cancer and prohormone		
XX	PT	processing		
XX	XX	Claim 1; Page 7; 26pp; English.		
XX	XX	This is the amino acid sequence of the human ASP2 (aspartic protease		
XX	CC	family), used in the method of the invention. Agonists and		
XX	CC	antagonists for ASP2 immunospecific antibodies are used to treat		
XX	CC	conditions requiring increased or decreased activity or expression of		
XX	CC	ASP2 respectively. ASP2 is used to treat and diagnose e.g.		
XX	CC	Alzheimer's disease, cancer and prohormone processing and ASP2 or a		
XX	CC	fragment can be used to induce an immune response against the above		
XX	CC	conditions.		
XX	XX	Sequence 501 AA;		
XX	Query Match	99.5%; Score 2650; DB 19; Length 501;		
XX	Best Local Similarity	99.6%; Pred. No. 1.5e-261;		
XX	Matches 499; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
XX	1	MAQALPWLMMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEPEEPPGRGSF	60	
XX	1	MAQALPWLMMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEPEEPPGRGSF	60	
XX	61	VEVDNLRKSGQGYVEMTVGSPQTLNLTVDGSSNFAVGAAPHPFLHRYRQLSST	120	
XX	61	VEVDNLRKSGQGYVEMTVGSPQTLNLTVDGSSNFAVGAAPHPFLHRYRQLSST	120	
XX	121	YDLRKGIVVPTQGWKEGELGTDLVSIHPGNVTVRANTAAITSDKFFINGSNNEGIL	180	
XX	121	YDLRKGIVVPTQGWKEGELGTDLVSIHPGNVTVRANTAAITSDKFFINGSNNEGIL	180	
XX	181	GLAYAEIAPDLSLEFFEDSLVKQTHVFNLFSLHLCGAGPLNGSEVLASVGGSMIIGGI	240	
XX	181	GLAYAEIAPDLSLEFFEDSLVKQTHVFNLFSLHLCGAGPLNGSEVLASVGGSMIIGGI	240	
XX	241	DHSLYTGSLWYTPIRREMYEYIIVRVEINGODLKMCKEYNYDKSIYDSTGNLRPKK	300	
XX	241	DHSLYTGSLWYTPIRREMYEYIIVRVEINGODLKMCKEYNYDKSIYDSTGNLRPKK	300	
XX	301	VFEAAVKSTKAASSTKEKPDGFWLGEOLVCMQAGTTPWNIIPVLSLYLMGEVTNQSPRIT	360	
XX	301	VFEAAVKSTKAASSTKEKPDGFWLGEOLVCMQAGTTPWNIIPVLSLYLMGEVTNQSPRIT	360	
XX	361	ILPQOYLRPVEDVATSDQCYKFAISQSSSTGTVMGAVIMEGFYVDFDRARRKIGFAVSAC	420	

PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
PA (UNII) UNIV ILLINOIS FOUND.
XX
PI Tang JJN, Hong L, Ghosh AK;

Fri Jun 6 08:30:58 2003

```

KW Memapsin 2; catalyst; Alzheimer's.
XX OS Homo sapiens.
XX PN WO200100663-A2.
XX PD 04-JAN-2001.
XX PF 27-JUN-2000; 2000WO-US17661.
XX PR 28-JUN-1999; 99US-0141363.
XX PR 30-NOV-1999; 99US-0168060.
XX PR 25-JAN-2000; 2000US-0177836.
XX PR 27-JAN-2000; 2000US-0178368.
XX PR 08-JUN-2000; 2000US-0210292.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PA Tang JUN, Lin X, Koelsch G;
XX WPI; 2001-102885/11.
XX PT Purified recombinant catalytically active memapsin 2, used to screen
XX PT inhibitors of it, which are used to treat and prevent Alzheimer's
XX PT disease -
XX PS
XX PS Claim 2; Page 73-75; 86pp; English.
XX CC The present invention relates to a purified recombinant
XX CC catalytically active memapsin 2. The invention may be used for
XX CC isolating inhibitors which are used to treat or prevent
XX CC Alzheimer's disease. The invention may also be used to screen
XX CC for individuals more genetically prone to develop Alzheimer's
XX CC disease.
XX SQ Sequence 488 AA;

Query Match 96.9%; Score 2582; DB 22; Length 488;
Best Local Similarity 99.8%; Pred. No. 1.3e-254;
Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSRFVEMVDNLRKSGQ 73
Db 1 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSRFVEMVDNLRKSGQ 60

QY 74 GYVEMTVGSPQTLNLLVDTGSSNFAVGAAPHLHRYYQRLSTYRDLRKGVVVPYT 133
61 GYVEMTVGSPQTLNLLVDTGSSNFAVGAAPHLHRYYQRLSTYRDLRKGVVVPYT 120

QY 134 QGKWEGLGTDLSIPHPNVTVRANTAAITSDKFFINGSNWEGILGLAYAEIARPPDS 193
121 QGKWEGLGTDLSIPHPNVTVRANTAAITSDKFFINGSNWEGILGLAYAEIARPPDS 180

QY 194 LEPPFDSLVRQTHVPNLFSLHLCAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLWYTP 253
181 LEPPFDSLVRQTHVPNLFSLHLCAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLWYTP 240

QY 254 IRREWYVEIIVRVEINGQDLKMDCKEYNDKSIYDVGTTNLRPKKVFEEAAVKSIAAS 313
241 IRREWYVEIIVRVEINGQDLKMDCKEYNDKSIYDVGTTNLRPKKVFEEAAVKSIAAS 300

QY 314 STEKFPDGFNLGDLVQWAGTTPWNIFPVISLYLMGEVTNQSFRTILPQOYLRPVEDV 373
301 STEKFPDGFNLGDLVQWAGTTPWNIFPVISLYLMGEVTNQSFRTILPQOYLRPVEDV 360

QY 374 ATSDDCKYKFAISQSSGTVMGAVIMEGFYVDFDRKRIGFAVSACHVHDEFRTAAVEG 433
361 ATSDDCKYKFAISQSSGTVMGAVIMEGFYVDFDRKRIGFAVSACHVHDEFRTAAVEG 420

QY 434 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICAFMLPLCLMVQCWRCLRCLRQHQHDF 493
421 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICAFMLPLCLMVQCWRCLRCLRQHQHDF 480
```

QY 494 ADDISLLK 501
Db 481 ADDISLLK 488

Search completed: June 5, 2003, 16:12:15
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:03:45 ; Search time 22 Seconds
(without alignments)
944.529 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLLMGAGVLPAP.....CLRLCROHDDFADDISLLK 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	2656	93.7	501	1 BACE_HUMAN	P56817 homo sapien
2	2569	96.4	501	1 BACE_RAT	P56819 rattus norv
3	2567	96.4	501	1 BACE_MOUSE	P56818 mus musculu
4	1173.5	44.1	518	1 BAE2_HUMAN	Q9Y520 homo sapien
5	330	12.4	324	1 PEPI_GADMO	P56272 gadus morhu
6	314.5	11.8	390	1 CATD_BOVIN	P80209 bos taurus
7	309	11.6	387	1 PEPI_RABIT	P28712 oryctolagus
8	307.5	11.5	388	1 PEPI_MACFU	P27678 macaca fusc
9	305	11.4	367	1 PEPA_CHICK	P00793 gallus gall
10	301.5	11.3	383	1 PEPE_CHICK	P16476 gallus gall
11	301.5	11.3	396	1 CATE_HUMAN	P14091 homo sapien
12	300.5	11.3	412	1 CATD_HUMAN	P07339 homo sapien
13	299	11.2	387	1 PEPI_RABIT	P27821 oryctolagus
14	298	11.2	387	1 PEPI_RABIT	P27821 oryctolagus
15	297	11.1	407	1 CATD_RAT	P24268 rattus norv
16	295	11.1	391	1 CATE_CAVPO	P25796 cavia porce
17	294.5	11.1	388	1 PEPI_MACFU	P27677 macaca fusc
18	289	10.8	387	1 PEPI_RABIT	P27822 oryctolagus
19	288.5	10.8	388	1 PEPA_HUMAN	P00790 homo sapien
20	288.5	10.8	388	1 PEPA_MACFU	P11489 macaca mula
21	288.5	10.8	398	1 CATE_RAT	P16228 rattus norv
22	287	10.8	410	1 CATD_MOUSE	P18242 mus musculu
23	286.5	10.8	388	1 PEPI_MACFU	P03954 macaca fusc
24	286	10.7	398	1 CATD_CHICK	Q05744 gallus gall
25	284.5	10.7	381	1 CHYM_SHEEP	P18276 ovis aries
26	281.5	10.6	386	1 PEPA_PIG	P00791 sus scrofa
27	281	10.5	387	1 PEPA_CALJA	Q9N2d4 callithrix
28	280.5	10.5	397	1 CATE_MOUSE	P70269 mus musculu
29	276.5	10.4	381	1 CHYM_BOVIN	P00794 bos taurus
30	276.5	10.4	396	1 CATE_RABIT	P43159 oryctolagus
31	274.5	10.3	419	1 CARV_CANAL	P10977 candida alb
32	273.5	10.3	376	1 PAG2_BOVIN	Q28057 bos taurus
33	273.5	10.3	377	1 PEPC_MACFU	P03955 macaca fusc

Result 1	ID	BACK_HUMAN	STANDARD	PRT	501 AA
AC	P56817	Q9UJT5; Q9BYC1; Q9BYC0; Q9BYB9;			
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Beta-secretase precursor (EC 3.4.23.-)	(Beta-site APP cleaving enzyme)			
DE	(Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl protease 2) (ASP 2) (ASP2) (Membrane-associated aspartic protease 2)				
DE	(Memapsin-2)				
GN	BACE OR BACE1				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM A).				
RC	TISSUE=Brain;				
RX	MEDLINE=20002972; PubMed=10531052				
RA	Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A., Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y., Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A., Biere A.L., Curran E., Burgess T., Louis J.C., Collins F., Treanor J., Rogers G., Citron M.;				
RT	"Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE."				
RL	Science 286:735-741(1999).				
[2]					
RP	SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND CHARACTERIZATION.				
RC	TISSUE=Brain;				
RX	MEDLINE=20057171; PubMed=10591214;				
RA	Sinha S., Anderson J.P., Barbour R., Basi G.S., Caccavellio R., Davis D., Doan M., Dovey H.F., Eriqon N., Hong J., Jacobson-Croak K., Jewett N., Keim P., Knops J., Lieberburg I., Power M., Tan H., Wang S., Tatsuno G., Tung J., Schenk D., Seubert P., Suomensaar S.M., Wang S., Walker D., Zhao J., McConlogue L., Varghese J.;				
RT	"Purification and cloning of amyloid precursor protein beta-secretase from human brain."				
RL	Nature 402:537-540(1999).				
[3]					
RP	SEQUENCE FROM N.A. (ISOFORM A).				
RX	MEDLINE=20057170; PubMed=10591213;				
RA	Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M., Brashier J.R., Strauman N.C., Mathews W.R., Buhl A.E., Carter D.B., Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;				
RT	"Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity."				
RL	Nature 402:533-537(1999).				
[4]					
RP	SEQUENCE FROM N.A. (ISOFORM A).				
RX	MEDLINE=20120043; PubMed=10656250;				
RA	Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D., Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M., Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;				

34	273	10.2	388	1	PEPF_RABIT	P27823 oryctolagus
35	270.5	10.2	381	1	CHYM_CALJA	Q9n2d2 callithrix
36	268	10.1	396	1	CARP_NEUCR	Q01294 neurospora
37	267	10.0	365	1	CATD_SHEEP	Q9nzs8 ovis aries
38	266.5	10.0	388	1	PEPC_CALJA	Q9n2d3 callithrix
39	266	10.0	394	1	PEPC_CAVPO	Q64411 cavia porce
40	266	10.0	405	1	CARP_YEAST	P02767 saccharomyc
41	264.5	9.9	388	1	PEPC_HUMAN	P20142 homo sapien
42	264	9.9	388	1	PAG_HORSE	Q28389 equus cabal
43	262	9.8	496	1	ASPP_ORYSA	P42211 oryza sativ
44	261.5	9.8	387	1	ASPP_AEDAE	Q03168 aedes aegypt
45	261	9.8	345	1	CATD_PIG	P00795 sus scrofa

ALIGNMENTS

Genew.	HCNC:933; BACE.
MM:	604252; -
DR	InterPro: IPR001461; AsproteaseA1.
DR	InterPro: IPR001969; Asprotease_site.
DR	Pfam: PF00026; asp: 1.
DR	PRINTS: PR00792; PEPsin.
DR	PROSITE: PS00141; ASP_PROTEASE; 1.
DR	PROSITE: PS00141; ASP_PROTEASE; 1.
KW	Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
KW	Signal; Alternative splicing, POTENTIAL.
FT	SIGNAL 1 21
FT	PROPEP 22 45
FT	CHAIN 46 501
FT	DOMAIN 22 457
FT	TRANSMEM 458 478
FT	DOMAIN 479 501
FT	DOMAIN 493 501
FT	ACT_SITE 93 93
FT	ACT_SITE 289 289
FT	DISULFID 216 420
FT	DISULFID 278 443
FT	DISULFID 330 380
FT	CARBOHYD 153 153
FT	CARBOHYD 172 172
FT	CARBOHYD 223 223
FT	CARBOHYD 354 354
FT	CARBOHYD 146 189
FT	VARSPLIC 190 214
FT	VARSPPLIC 190 214
SQ	SEQUENCE 501 AA: 55763 MW: 377C64C824ACEF05 CRC64;
	Query Match 99.7%; Score 2656; DB 1; Length 501;
	Best Local Similarity 99.8%; Pred. No. 8.3e-206;
	Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 MAQALPWLIIWMGAGVLPAGHTQHGIRLPURSLGGAGPLGLRLPRETDEEPEEGRGGSF 60
DB	1 MAQALPWLIIWMGAGVLPAGHTQHGIRLPURSLGGAGPLGLRLPRETDEEPEEGRGGSF 60
QY	61 VEMVDNLRKSGQGYVEMTVGSPQPTLNLVDVTGSSNFVAGAAHPFLHRYQRLSST 120
DB	61 VEMVDNLRKSGQGYVEMTVGSPQPTLNLVDVTGSSNFVAGAAHPFLHRYQRLSST 120
QY	121 YRDLRKGVYVPTQCKWEGELGTLDSIPHGPNVTYRANIAAITSDEKFFINGSNWEGIL 180
DB	121 YRDLRKGVYVPTQCKWEGELGTLDSIPHGPNVTYRANIAAITSDEKFFINGSNWEGIL 180
QY	181 GLAYAEIARPDSDLPEFFDSLVKQTHVPNPLFSLHLCGAGPPLNQSEVLASVCGSMIIGGI 240
DB	181 GLAYAEIARPDSDLPEFFDSLVKQTHVPNPLFSLQLCGAGPPLNQSEVLASVCGSMIIGGI 240
QY	241 DLSLYTGSWLYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNLRPKK 300
DB	241 DLSLYTGSWLYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNLRPKK 300
QY	301 VFEAAVKSIKAASSTKEKPDGFWLGEQLVCWQAGTTPWNTFPVLSLYLMEGVTNQSFRT 360
DB	301 VFEAAVKSIKAASSTKEKPDGFWLGEQLVCWQAGTTPWNTFPVLSLYLMEGVTNQSFRT 360
QY	361 ILPQOYLRPVEDVATSSQDDCKYFALISQSSSTGVVNGAVIMEGFYVDFDRARKRIGFAVSAC 420
DB	361 ILPQOYLRPVEDVATSSQDDCKYFALISQSSSTGVVNGAVIMEGFYVDFDRARKRIGFAVSAC 420
QY	421 HVHDEFTFAVEGPFVTLDMEDCGYNIPQDDESTLMTIATYMAAICALFMLPLCLMVCOW 480
DB	421 HVHDEFTFAVEGPFVTLDMEDCGYNIPQDDESTLMTIATYMAAICALFMLPLCLMVCOW 480
QY	481 RCLRLRQOHHDDFADDISLLK 501
DB	481 RCLRLRQOHHDDFADDISLLK 501

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
DE protease 2) (ASP2) (Membrane-associated aspartic protease 2)
DE (Memapsin-2).
GN BACE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20002972; PubMed=10531052;
RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A., Luo Y.,
RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
RA Fisher S., Fuller J., Edenson S., Lille J., Jarosinski M.A.,
RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
RA Treanor J., Rogers G., Citron M.;
RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RL the transmembrane aspartic protease BACE.";
RL Science 286:735-741(1999).
CC -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
CC AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
CC THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
CC SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
CC FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF190727; AAF04144.1; -
CC HSPB; P33229; IYPS.
CC MERO; A01.004; -
CC InterPro: IPR001461; AsparticaseA1.
CC InterPro: IPR001969; AsparticaseA1.
CC Pfam: PF00026; asp. 1.
CC PRINTS: PF00792; PEPsin.
CC PROSITE: PS00141; ASP_PROTEASE; 1.
KW Hydrolyase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 45 POTENTIAL.
FT CHAIN 46 501 BETA-SECRETASE.
FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 458 478 POTENTIAL.
FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 289 289 BY SIMILARITY.
FT DISULFID 216 420 BY SIMILARITY.
FT DISULFID 278 443 BY SIMILARITY.
FT DISULFID 330 380 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 501 AA; 55806 MW; 24B445BC8BE87DE3 CRC64;
Query Match 96.4%; Score 2569; DB 1; Length 501;
Best Local Similarity 96.2%; Pred. No. 8.2e-199;
Matches 482; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
Qy 1 MAQALPWLMMGAGVLPAGHTQHGIRLPLRSLGGAPLGLRLPRETDEEPEPGRGSF 60

Db 1 MAPALRWLLLVGSGMLPAQOGLHGLRLPLRSLGAPLGLRLPRETDEEPEPGRGSF 60
Qy 61 VEMVDNLRGSGGQGYVEMTVGSPPTQNLILVDGSSNFVAGAAPHPFLHRYQRLSST 120
Db 61 VEMVDNLRGSGGQGYVEMTVGSPPTQNLILVDGSSNFVAGAAPHPFLHRYQRLSST 120
Qy 121 YRDLRKGVYYPYTOGKWEGLGTLVSTPHGPNVTVRANITAAITSDKFFINGSNWEGIL 180
Db 121 YRDLRKSVYYPYTOGKWEGLGTLVSTPHGPNVTVRANITAAITSDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLPEPFDSLVKQTHVPNLFSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDLPEPFDSLVKQTHVPNLFSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGLSMYTPIRREMYVEIIVRVEINQDLKMDCKEYNDKSIDVSGTTNLRPKK 300
Db 241 DHSLYTGLSMYTPIRREMYVEIIVRVEINQDLKMDCKEYNDKSIDVSGTTNLRPKK 300
Qy 301 VFEEAAVKSIRAASTEFKPDGFWLGEOLVQWAGCTTPWNPFPVLSLYLMGEVTVNQSFRT 360
Db 301 VFEEAAVKSIRAASTEFKPDGFWLGEOLVQWAGCTTPWNPFPVLSLYLMGEVTVNQSFRT 360
Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGFYVDFRARRKRGFVAVSAC 420
Db 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGFYVDFRARRKRGFVAVSAC 420
Qy 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQVDESLMTIAYVMAAICAFMLPCLMVCOW 480
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQVDESLMTIAYVMAAICAFMLPCLMVCOW 480
Qy 481 RCLRLRQHQHDDFADDSLLK 501
Db 481 RCLRLRQHQHDDFADDSLLK 501
RESULT 3
ID BACE_MOUSE STANDARD; PRY; 501 AA.
AC P56818;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
DE protease 2) (ASP2) (Membrane-associated aspartic protease 2)
DE (Memapsin-2).
GN BACE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20002972; PubMed=10531052;
RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
RA Fisher S., Fuller J., Edenson S., Lille J., Jarosinski M.A.,
RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
RA Treanor J., Rogers G., Citron M.;
RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RT the transmembrane aspartic protease BACE.";
RL Science 286:735-741(1999).
RN [2]
RN REVISIONS TO 6 AND 81-87.
RA Bennett B.D., Vassar R., Citron M.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20051710; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA Brashler J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA Tomasselli A.G., Parodi L.A., Heinrichson R.L., Gurney M.E.;

RN [3] SEQUENCE FROM N.A.
 RP Acciarino M.P., Fumagalli P., Ottolenghi S., Taramelli R.;
 RA "Cloning of a gene from chromosome 21 Down region encoding a potential
 RT transmembrane aspartyl protease";
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [4] SEQUENCE FROM N.A.
 RP Solans A., Estivill X., de la Luna S.;
 RA "Cloning of a novel mammalian aspartyl protease";
 RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RL [5]
 RN [5] SEQUENCE FROM N.A.
 RP MEDLINE=20120043; PubMed=10656250;
 RX Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
 RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
 RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
 RT "Identification of a novel aspartic proteinase (Asp 2) as
 RL beta-secretase";
 RL Mol. Cell. Neurosci. 14:419-427(1999).
 RN [6] SEQUENCE FROM N.A.
 RP


```

Query Match      12.4%; Score 330; DB 1; Length 324;
Best Local Similarity 27.9%; Pred. No. 3.7e-19;
Matches 104; Conservative 67; Mismatches 136; Indels 66; Gaps 15;

QY 63 MVDNLKSGQGYVEMTVGSPPTLNILVDTGSSNFVAG----AAPHFLHRYVORLS 118
DB 2 VTEQMKNEADTEYGVISIGTTPESFKVIFDTGSSNLWVSSSHCSAQACSNHNKFKPQS 61

QY 119 STYRDLRKGVYDYTGKWEGLGTLVSIPIHG--PNYTVRANIAAITESDKFEINGSNW 176
DB 62 STYVETGKTVDLRYTGCGMRGILGQDVTSGVGGSPNQELG---ESQTEPGFQA-AAPF 117

QY 177 EGIILGLAYARIAPDSDLEPFDSLKVQTHV--PNLFLSLHLCAGFPPLNOSSEVLASVGGSM 235
DB 118 DGILGLAYPSIAAA--GAYPVFDNMGSSQILVEKDLFFSYLGGG--ANGSEVM----- 166

QY 236 IIGIDHSLVTLGSLWTPYPIRREYVYVIVRVEINGQDLKMD--CKEYNDKSIDVSGTGN 294
DB 167 -LGVDSNHYTGSHTWIPVTAERYQWALDITVNGQTAAECG-----QALVDTGTSK 219

QY 295 LRLPKKVFEEAAVKSIAASSTKFPDGLVQVQAGTTPWNIFFPVISLYLMGEVTN 354
DB 220 IVAPVSALANIMKDIGASEN-----OGEMGN--CASVQSLPDITF-----TI 260

QY 355 QSFRTITLPOOYLRPVEDVATSDCKYKFAISOSSSTGT-----VMGAVIMEGFYVVF 406
DB 261 NGVKOPLPPSAYIEGQAFCTSD-----GLGSGVPSNTSELWIFGDVFLURNYITTY 311

QY 407 DRARKRIGFAVSA 419
DB 312 DRTNKKVGFAPAA 324

RESULT 6
CATD_BOVIN STANDARD; PRT; 390 AA.
ID AC CATD_BOVIN STANDARD; PRT; 390 AA.
AC P80209; Q9TS27;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
GN CTSD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
SEQUENCE OF 1-48.
TISSUE=Milk;
MEDLINE=93202276; PubMed=8454061;
Larsen L.B., Boisen A., Petersen T.E.;
"Procathepsin D cannot autoactivate to cathepsin D at acid pH.";
FEBS Lett. 319:54-58(1993).
[2]
SEQUENCE OF 45-390, AND X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93223670; PubMed=8467789;
RA Metcalf P., Fusek M.;
"Two crystal structures for cathepsin D: the lysosomal targeting
signal and active site.";
EMBO J. 12:1293-1302(1993).
RL ENBO J. 12:1293-1302(1993).
CC -!- FUNCTION: Acid protease active in intracellular protein breakdown.
CC -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC that of pepsin A. Does not cleave the 4-Gln-His-5 bond in B
CC chain of insulin.
CC -!- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR PIR: S32383; S32383.
DR PIR: S37419; S37419.
DR HSSP: P07339; 1LYB.

```

```

MEROPS; A01.009;
DR InterPro: IPR001461; AsparticaseA1.
DR InterPro: IPR001969; Asparticase_site.
DR Pfam: PF00026; asp; 1.
DR PRINTS: PR00792; PEPSIN.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Zymogen.
FT PROPEP 1 44
FT CHAIN 45 390
FT ACT_SITE 77 77
FT ACT_SITE 273 273
FT DISULFID 71 140
FT DISULFID 90 97
FT DISULFID 264 268
FT DISULFID 307 344
FT CARBOHYD 114 114
FT CARBOHYD 241 241
SQ SEQUENCE 390 AA; 4248 MW; 5B38AAIC33C48D35 CRC64;

Query Match      11.8%; Score 314.5; DB 1; Length 390;
Best Local Similarity 28.0%; Pred. No. 8.4e-18;
Matches 113; Conservative 72; Mismatches 128; Indels 91; Gaps 21;

QY 53 EPG-RRGSFVEMVDNLRCKSQGYVEMTVGSPPTLNILVDTGSSNFVAGAAPHPFL-- 109
DB 39 EPAVRGQPIPELLKNYMDAQ---YGEIGIGTTPQCFVVDFTGSANLWVPSIHCLLDI 95

QY 110 ---HRYQRLSSSTYRDLRKGVY--VPYTGKWEGLGTLVSIPIHGPN-----VTVR 157
DB 96 ACWTHRYKNSDKSTY--VKNGTTFDIHYGSGSLGYSQDTSVPCNPSSSPGGVTVO 153

QY 158 ANI--AAITSDKTFINGSNWEGILGLAYAEIARPDSDLEPFDSLKVQTHV--PNLFLSL 214
DB 154 RQTGGEAIKQGVVFI-AAKFDGILGMAYPRIS--VNNLVDFVFNLMQOKLVKNVFS-- 208

QY 215 LCGAGFPLNOSSEVLASVCGSMIIGIDHSLVTLGSLWTPYPIRREYVYVIVRVEINGQDL 274
DB 209 ---FFLNR-DPKAOPGGLMGLGTDSDSKYRGSLMHNVTROAYWQTHMDQLDV-GSSL 261

QY 275 KMDCKEYNYDKSIDVSGTNTNLRPKKVFEEAAVKSIAASSTKFPDGLVQVLF-CWQA 333
DB 262 TV-CK--GGCEAIVDTGTSIVGPEVEERLQKAGVPLIQ-----GEYMIPEKV 310

QY 334 GTTPWNIFFVISLYLMGEVTNQSFRITLPOOYLRPVEDVATSDCKYKFAISOSSSTGT- 392
DB 311 SS-----LPEVIVKLG-----KDYALSPED-YALKVSOAETVVC 344

QY 393 -----VMGAVIMEGFYVVFDRARKRIGFAVSA 419
DB 345 LSGFMGMDIPPPGGLWILGDFVIGRYTVFDRDQNRVGLAEAA 388

RESULT 7
PEPI_RABIT STANDARD; PRT; 387 AA.
ID PEPI_RABIT STANDARD; PRT; 387 AA.
AC P28712;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pepsin II-1 precursor (EC 3.4.23.1) (Pepsin A).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=91009127; PubMed=2129536;
RX Kageyama T., Tanabe K., Koilwai O.;
"Structure and development of rabbit pepsinogens. Stage-specific
zymogens, nucleotide sequences of cDNAs, molecular evolution, and
gene expression during development.";
RL J. Biol. Chem. 265:17031-17038(1990).
CC -!- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS

```


Matches	88; Conservative	70; Mismatches	124; Indels	84; Gaps	13;
QY	75	YVEMTVGSPQTLNLIIVDTGSSNFVGAAPHPFL-----HRYTORQLSSTYRDLRKG	127		
Db	59	YGTITISCTPOQDSVIFDTGSSNLWV---PSIYCKSSACSNHKRFPDPSKSTVYSTNET	115		
QY	128	VVVPYTGQKGEGLGTLVSIHPGNVTVRANTAAATESDK--FFINGSNWEGILGLAYAE	186		
Db	116	VVIAYGTGSMGILGYDITVAV---SSIDVQNOIFGLSETEPGFFYYCNFDGILGLAFPS	172		
QY	187	IARPPDLSLEPPFDSLVKQTHV-PNLFSLHLCGAGFPLNQSEVLASVGGSMITGGIDHSLY	245		
Db	173	IS--SSGATPVFDNINWNRSLVSQDLFSVLSKDG-----ETGSFVLFGIDPNYT	220		
QY	246	TGSLWMTPIRREWEYVEIIVRVEINGQDLK--MDCKEYNDKSIYDSGTTNLRPKKVEE	303		
Db	221	TGKIYVPLPSAETVWQIITMDRVTGNKYVACFFTC-----QAIVDTGTSLLVWPOGAYN	274		
QY	304	AAVKSIAKASSTE-----KFPDGFGLGEQLVCWQAGTTPWNIFFVISLYLMEVYNSQ	356		
Db	275	RIIKDLGVSSDGEISCDISKLPD-----VTEHINGHA-----	307		
QY	357	FRITILPOQYLRPVEDVATSDQCYKFAISOSSTGT-----VMGAVIMEGFYVDFDRAR	410		
Db	308	-----FTLPASAYVLNEDGSCMLGFENMGTPTELGEQWILGDVFIQYFTVFDRAN	358		
QY	411	KRIGFA	416		
Db	359	NKVGLS	364		
RESULT 10					
PEPE_CHICK					
ID	PEPE_CHICK	STANDARD;	PRT;	383	AA.
AC	P16476;				
DT	01-AUG-1990	(Rel. 15, Created)			
DT	01-AUG-1990	(Rel. 15, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Embryonic pepsinogen precursor (EC 3.4.23.-).				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
OX	NCBI_TaxID=9031;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88227903;	PubMed=3131317;			
RA	Hayashi K., Agata K., Mochii M., Yasugi S., Eguchi G., Mizuno T.;				
RT	"Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsinogen: phylogenetic relationship with prothymosin.";				
RL	J. Biochem. 103:290-296(1988).				
CC	-!- DEVELOPMENTAL STAGE: SPECIFICALLY SECRETED DURING THE EMBRYONIC PERIOD IN THE CHICKEN PROVENTRICULUS (GLANDULAR STOMACH).				
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; D00215; BAA00153.1; ..				
DR	PIR; A41443;	4CMS.			
DR	HSSP; P00794;	4CMS.			
DR	MEROPS; A01.028;				
DR	InterPro; IPR001461;	Asparticase1.			
DR	InterPro; IPR001969;	Asparticase_site.			
DR	Pfam; PF00026;	asp; 1.			
DR	PRINTS; PR00792;	PEPSIN.			
DR	PROSITE; PS00141;	ASP_PROTEASE; 2.			
KW	Hydrolase; Aspartyl	protease; Digestion; Zymogen; Glycoprotein.			
FT	PROPEP	1	42		
FT	CHAIN	43	367		
FT	ACT_SITE	77	77		
FT	ACT_SITE	260	260		
FT	CARBOHYD	113	113	N-LINKED (GLCNAC. . .).	
FT	DISULFID	90	95		
FT	DISULFID	251	255		
FT	DISULFID	290	323		
SQ	SEQUENCE	367	AA;	40431	MW; OC547E7FD8F5B341 CRC64;
Query Match					
Best Local Similarity					
11.4%; Score 305; DB 1; Length 367;					
24.0%; Pred. No. 4.5e-17;					

RESULT 10

PEPE_CHICK

ID PEPE_CHICK STANDARD; PRT; 383 AA.

AC P16476;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Embryonic pepsinogen precursor (EC 3.4.23.-)

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88227903; PubMed=3131317;

RA Hayashi K., Agata K., Mochii M., Yasugi S., Eguchi G., Mizuno T.;

RT "Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsinogen: phylogenetic relationship with prochymosin.";

RL J. Biochem. 103:290-296(1988).

CC -!- DEVELOPMENTAL STAGE: SPECIFICALLY SECRETED DURING THE EMBRYONIC PERIOD IN THE CHICKEN PROVENTRICULUS (GLANDULAR STOMACH).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D00215; BAA00153.1;

DR PIR; A41443; A41443.

DR HSSP; P00794; 4CMS.

DR MEROPS; A01.028;

DR InterPro; IPR001461; Asparticase1.

DR InterPro; IPR001969; Asparticase_site.

DR Pfam; PF00026; asp; 1.

DR PRINTS; PR00792; PEPSIN.

DR PROSITE; PS00141; ASP_PROTEASE; 2.

DR Hydrolase; Aspartyl

KW

[illegible]

```

RESULT 11
CATE_HUMAN
ID CATE_HUMAN STANDARD; PERT; 396 AA.
AC P14091;
DE 01-JAN-1990 (Rel. 13, Created)
D1 01-JAN-1990 (Rel. 13, Last sequence update)
D2 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin E precursor (EC 3.4.23.34).
CTSE.
GN GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89380302; PubMed=2674141;
RA Azuma T., Pals G., Mohandas T.K., Couvreur J.M., Taggart R.T.;
RT "Human gastric cathepsin E. Predicted sequence, localization to
RT chromosome 1, and sequence homology with other aspartic
RT proteinases.";
RL J. Biol. Chem. 264:16748-16753(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112877; PubMed=1370478;
RA Azuma T., Liu W.G., Vander Laan D.J., Bowcock A.M., Taggart R.T.;
RT "Human gastric cathepsin E gene. Multiple transcripts result from
RT alternative polyadenylation of the primary transcripts of a single
RT gene locus at 1q31-q32.";
RL J. Biol. Chem. 267:1609-1614(1992).
RN [3]

```

```

SEQUENCE FROM N.A.
RP Tatnell P.J., Kay J.;
RA "Human cathepsin E.";
RT Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN [4]
RL
RR
RS
RW
RX SEQUENCE OF 54-68; 77-95; 141-155; 275-285 AND 389-396.
RY MEDLINE=90241267; PubMed=2334440;
RZ Athauda S.B.P., Matsuzaki O., Kageyama T., Takahashi K.;
SA Structural evidence for two isozymic forms and the carbohydrate
SB attachment site of human gastric cathepsin E.";
SC Biochem. Biophys. Res. Commun. 168:878-885(1990).
SD CC -!- FUNCTION: DUE OT ITS INTRACELLULAR LOCATION AND DISTRIBUTION IN
SE LYMPHOID ASSOCIATED TISSUE, IT MAY HAVE A ROLE IN IMMUNE FUNCTION.
SF CC -!- CATALYTIC ACTIVITY: Similar to cathepsin D, but slightly broader
SG specificity.
SH CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
SI CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
SJ CC -----
SK This SWISS-PROT entry is copyright. It is produced through a collaboration
SL between the Swiss Institute of Bioinformatics and the EMBL Outstation -
SM the European Bioinformatics Institute. There are no restrictions on its
SN use by non-profit institutions as long as its content is in no way
SO modified and this statement is not removed. Usage by and for commercial
SP entities requires a license agreement (See http://www.isb-sib.ch/announce/
SQ or send an email to licenses@isb-sib.ch).
SR
SS EMBL; M84424; AAA52300.1; -.
ST EMBL; M84413; AAA52300.1; JOINED.
SU EMBL; M84417; AAA52300.1; JOINED.
SV EMBL; M84418; AAA52300.1; JOINED.
SW EMBL; M84419; AAA52300.1; JOINED.
SX EMBL; M84420; AAA52300.1; JOINED.
SY EMBL; M84421; AAA52300.1; JOINED.
SZ EMBL; M84422; AAA52300.1; JOINED.
TA EMBL; J05036; AAA52130.1; -.
TB EMBL; AJ250717; CAB82850.1; -.
TC FIR; A34401; A34401.
TD FIR; A34643; A34643.
TE FIR; A42038; A42038.
TF HSSP; P00794; 4CMS.
TG MEROPS; A01.010; -.
TH Genew; HGNC:2530; CTSE.
TI MIM; 116890;
TJ InterPro; IPR001461; AspproteaseA1.
TK InterPro; IPR001969; Aspserine_sit_e.
TL Pfam; PF00026; asp; 1.
TM PRINTS; PR00792; PEPSTN.
TN PROSITE; PS00141; ASP_PROTEASE; 2.
TO Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Signal.
TP SIGNAL 1 17
TT PROPEP 18 53 ACTIVATION PEPTIDE.
TU CHAIN 54 396 CATHEPSIN E.
TV MOD_RES 18 18 PYRROLIDONE CARBOXYLIC ACID.
TW ACT_SITE 96 96 BY SIMILARITY.
TX ACT_SITE 281 281 BY SIMILARITY.
TY DISULFID 60 60 INTERCHAIN (PROBABLE).
TZ DISULFID 109 114 BY SIMILARITY.
UA DISULFID 272 276 BY SIMILARITY.
UB DISULFID 314 351 BY SIMILARITY.
UC CARBOHYD 90 90 N-LINKED (GLCNAC... ).
UD CARBOHYD 220 220 O-LINKED (POTENTIAL).
UE CARBOHYD 333 333 O-LINKED (POTENTIAL).
UF SEQUENCE 396 AA; 42793 MW; 40B643C5FB01521E CRC64;

```

```

Query Match      11.3%  Score 301.5;  DB 1;  Length 396;
Best Local Similarity 25.8%;  Pred. No. 9.5e-17;
Matches 100;  Conservative 68;  Mismatches 144;  Indels 75;  Gaps 16;

QY 48 DEEPEPGRRGSGFVEMVDNLKRGSGGYVEMTVGSPQDTNLIVDTGSSNFAVGA---- 103
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 63 DSAKEP-----LINVLD-----MEYFETISGSPQNFTVFTDGSNLMWPSVYCT 110
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 104 AHPFHLRYHQROLSSYTRDLRKGVVVPVYTGKWEGLGTDLYSIPIHGPNVTVRANIAAI 163

```

111 SPACKTHSRQPSQSTYSQPGSQSFIQYQSGLSGIGADQVSV-EGLTVVQGOQFGEV 169
164 TESDKFFINGSNNEGILGLAYAEIARPDSDLEFFDSLVKQTHVP-NLFSHLGCGAFPL 222
170 TEPQGTQVD-AEFDGILGLGYPSLA--VGGVTPVFDNMMNAQNVLDPMSFVYM----- 219
223 NOSEVLASVSGSMIGIDHSLVTSGLWYTPIRREWYEVIIYVRVEINGODLKMOCKEYN 282
220 -SSNPEGAGSELIFGCDYHSHFSGSLNWPVTKQAYWQIALDNIQVGG--TVMFCSE-- 274
283 YDKSIVSGTNNLRPKKVFEEAAVKSKAASSTKEKFPDGFNLGQVLCWQAGTTPWNIPP 342
275 GCOAIVDTGTSITGCPDKIKQLQNAIGAAP-----VDGEYAVE-----CANLVMP 321
343 VISLYLMGEVNTQSFRTILPQOVLRPVEDVATSDQDCYKFAISQSTG----- 391
322 DVTFITNG-----VPYTLSTPAY--TLLDFVDMQFEC-----SSGFGQLDIHPAG 365
392 --TVMGAVIMEGYVVFEDRARKRIGFA 416
366 PLWILGDFVIRQYFVDFGRNVRVGLA 392

RESULT 12
CATD_HUMAN
ID CATD_HUMAN STANDARD; PRT; 412 AA.
AC P07339;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
GN CTSD.
OS "Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270436; PubMed=3927292;
RA Faust P.L., Kornfeld S., Chirgwin J.M.;
RT "Cloning and sequence analysis of cDNA for human cathepsin D.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87231058; PubMed=3598310;
RA Westley B.R., May F.E.B.;
RT "Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive
human breast cancer cells.";
Nucleic Acids Res. 15:3773-3786(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91299159; PubMed=2069717;
RA Redeker B., Heckendorf B., Grosch H.W., Mersmann G., Hasilik A.;
RT "Molecular organization of the human cathepsin D gene.";
RL DNA Cell Biol. 10:423-431(1991).
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE=94085791; PubMed=8262386;
RA May F.E., Smith D.J., Westley B.R.;
RT "The human cathepsin D-encoding gene is transcribed from an estrogen-
regulated and a constitutive start point.";
RL Gene 134:277-282(1993).
RN [6]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE=95021301; PubMed=7935485;
RA Augereau P., Miralles F., Cavailles V., Gaudelet C., Parker M.,
RA Rochefort H.;

"Characterization of the proximal estrogen-responsive element of
human cathepsin D gene.";
Mol. Endocrinol. 8:693-703(1994).
RN [7]
RP SEQUENCE OF 170-180.
RC TISSUE=Liver;
RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargass R.,
RA Appel R.D., Hughes G.J.;
RL Submitted (JUN-1992) to the SWISS-PROT data bank.
RN [8]
RP VARIANT VAL-58.
RX MEDLINE=20179010; PubMed=10716266;
RA Papassotiropoulos A., Bagli M., Kurz A., Kornhuber J., Forstl H.,
RA Maier W., Pauls J., Lautenschlager N., Heun R.;
RT "A genetic variation of cathepsin D is a major risk factor for
Alzheimer's disease.";
RL Ann. Neurol. 47:399-403(2000).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RC TISSUE=Spleen;
RX MEDLINE=93223670; PubMed=8467789;
RA Metcalf P., Fusek M.;
RT "Two crystal structures for cathepsin D: the lysosomal targeting
signal and active site.";
RL EMBO J. 12:1293-1302(1993).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93342076; PubMed=8393577;
RA Baldwin E.T., Bhat T.N., Gulnik S., Hosur M.V., Sowder R.C. II,
RA Cachau R.E., Collins J., Silva A.M., Erickson J.W.;
RT "Crystal structures of native and inhibited forms of human cathepsin
D: implications for lysosomal targeting and drug design.";
Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).
RN [11]
RP FUNCTION: Acid protease active in intracellular protein breakdown.
CC Involved in the pathogenesis of several diseases such as breast
cancer and possibly Alzheimer's disease.
CC -I- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
that of pepsin A. Does not cleave the 4-Gln-1-His-5 bond in B
chain of insulin.
CC -I- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC -I- SUBCELLULAR LOCATION: Lysosomal.
CC -I- POLYMORPHISM: The Val-58 allele is significantly overrepresented
in demented patients (11.8%) compared with nondemented controls
(4.9%). Carriers of the Val-58 allele have a 3.1-fold increased
risk for developing AD than noncarriers.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M11233; AAB59529.1; -;
CC EMBL; X05344; CA28955.1; -;
CC EMBL; M63138; AAA51922.1; -;
CC EMBL; M63134; AAA51922.1; JOINED.
CC EMBL; M63135; AAA51922.1; JOINED.
CC EMBL; M63136; AAA51922.1; JOINED.
CC EMBL; M63137; AAA51922.1; JOINED.
CC EMBL; BC016320; AAI16320.1; -;
CC EMBL; L12980; AAI16314.1; -;
CC EMBL; S74899; AAI1456.1; -;
CC EMBL; S52557; AAI13668.1; -;
CC PIR; A25771; KKHUO.
CC PDB; 1LYA; 31-JAN-94.
CC PDB; 1LYB; 31-JAN-94.
CC MEROPS; A01.009; -;
CC SWISS-2DPAGE; P07339; HUMAN.

Fri Jun 6 08:31:00 2003

```

CC CC      -!- DEVELOPMENTAL STAGE: PEPSINOGENS IN GROUP I, II, AND III WHERE
CC CC      THE PREDOMINANT ZYMogens AT LATE POSTNATAL STAGE.
CC CC      -!- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY
CC CC      HORMONES AND RELATED SUBSTANCES.
CC CC      -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC CC
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC
CC CC      EMBL: M59235; AA85369.1; -
CC CC      PIR: C38302; C38302.
CC CC      HSP: P00790; IPSN.
CC CC      MEROPS: A01.001; -.
CC CC      InterPro: IPR001461; AspproteaseA1.
CC CC      InterPro: IPR001969; Aspprotease_site.
CC CC      Pfam: PF00026; asp; 1.
CC CC      PRINTS: PR00792; PEPsin.
CC CC      PROSITE: PS00141; ASP_PROTEASE; 2.
CC CC      KW Hydrolyase; Aspartyl protease; Digestion; Zymogen; Signal;
CC CC      Phosphorylation; Multigene family.
CC CC      SIGNAL 1 15
CC CC      FT PROPEP 16 59 ACTIVATION PEPTIDE.
CC CC      FT CHAIN 60 387 PEPsin II-2/3.
CC CC      FT MOD_RES 129 129 PHOSPHORYLATION (POTENTIAL).
CC CC      FT ACT_SITE 93 93 BY SIMILARITY.
CC CC      FT ACT_SITE 276 276 BY SIMILARITY.
CC CC      FT DISULFID 106 111 BY SIMILARITY.
CC CC      FT DISULFID 267 271 BY SIMILARITY.
CC CC      FT DISULFID 310 343 BY SIMILARITY.
CC CC      FT SEQUENCE 387 AA; 42100 MW; 66FC31A3DC75891 CRC64;
CC CC
CC CC      Query Match 11.2%; Score 299; DB 1; Length 387;
CC CC      Best Local Similarity 26.9%; Pred. No. 1.5e-16;
CC CC      Matches 97; Conservative 64; Mismatches 134; Indels 66; Gaps 13;
CC CC
CC CC      75 YVEMTVGSPQPTNLIVDTGSSNFVGAAPHF-----LHRYQRLSSYRDLRKG 127
CC CC      75 YFTGISICTPQDFTVIFDTGSSNLW---PSTYCSLALCHKRFPEDSSYTGTS 131
CC CC
CC CC      128 VYVPTQCKWEGELGTLVSIPIHGNPNTVRANIAAITESDKFFINGSGILGAYAEI 187
CC CC      132 LSITYGTGSMTGILGYDITVVKVGSIEDTNIQIFGLSKTEPLTF 189
CC CC
CC CC      188 ARPDLSLEPFDFSLVKQTHV-PNLFSLHLCGAGFPLNQSEVLASVGGSMIIIGIDHSLYT 246
CC CC      190 SSSDAT--PVFDNMWNEGLVSQDLFSVYLSDD-----EKGSLVMFGIDSSYTT 237
CC CC
CC CC      247 GSLWYTPIRREWYEVIIIVRVEINGDLKM--DCKEYNDKSIDVDSGTTNLRPKKVEFA 304
CC CC      238 GSLNWPVSYEGYQWITMDSVINGETIACADSC-----QAVDGTGSLTGP-----TS 287
CC CC
CC CC      305 AVKSKTAASTKEKFPDGFGLGEOLV-CWQAGTTPWNIFPVISLYLGEVNTQSFRTILP 363
CC CC      288 AISNIQSYIGASK-----NLLGENVISCAIDSLPDIVF-----TING 325
CC CC
CC CC      364 QOYLRPVEDVATSDQCKYFAISOSTGT-----VMGAVIMEGFYVVDPRARKTGFAV 417
CC CC      326 IQYPLPASAYILKEDDCTSGLEGNMVDTYTGLWILGDVFIRQYFTVDRANNQUGLAA 385
CC CC
CC CC      418 S 418
CC CC      386 A 386
CC CC
CC CC      RESULT 14
CC CC      ID PEP4_RABIT
CC CC      AC P28713;

```

```

DT DT      01-DEC-1992 (Rel. 24, Created)
DT DT      01-DEC-1992 (Rel. 24, Last sequence update)
DT DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE DE      Pepsin II-4 precursor (EC 3.4.23.1) (Pepsin A).
OS OS      Oryctolagus cuniculus (Rabbit).
OC OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC OC      Mammalia; Eutheria; Lagomorpha; Leporidae;
OC OC      NCBI_TaxID=9986;
RN RN      SEQUENCE FROM N.A.
RX RX      MEDLINE=91009127; PubMed=2129536;
RA RA      Kageyama T., Tanabe K., Koizumi O.;
RT RT      "Structure and development of rabbit pepsinogens. Stage-specific
RT RT      zymogens, nucleotide sequences of cDNAs, molecular evolution, and
RT RT      gene expression during development."
RL RL      J. Biol. Chem. 265:17031-17038(1990).
CC CC      -!- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
CC CC      INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC CC      ALSO CLEAVED TO SOME EXTENT.
CC CC      -!- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably
CC CC      aromatic, residues in pi and pi' positions. Cleaves 1-Phe-I-Val-2,
CC CC      4-Gln-I-His-5, 13-Glu-I-Ala-14, 14-Ala-I-Leu-15, 15-Leu-I-Tyr-16,
CC CC      16-Tyr-I-Leu-17, 23-Gly-I-Phe-24, 24-Phe-I-Phe-25 and 25-Phe-I-
CC CC      Tyr-26 bonds in the B chain of insulin.
CC CC      -!- DEVELOPMENTAL STAGE: PEPSINOGENS IN GROUP I, II, AND III WHERE
CC CC      THE PREDOMINANT ZYMogens AT LATE POSTNATAL STAGE.
CC CC      -!- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY
CC CC      HORMONES AND RELATED SUBSTANCES.
CC CC      -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC CC      PIR: D38302; D38302.
CC CC      HSP: P00790; IPSN.
CC CC      MEROPS: A01.001; -.
CC CC      InterPro: IPR001461; AspproteaseA1.
CC CC      InterPro: IPR001969; Aspprotease_site.
CC CC      Pfam: PF00026; asp; 1.
CC CC      PRINTS: PR00792; PEPsin.
CC CC      PROSITE: PS00141; ASP_PROTEASE; 2.
CC CC      KW Hydrolyase; Aspartyl protease; Digestion; Zymogen; Signal;
CC CC      Phosphorylation; Multigene family.
CC CC      SIGNAL 1 15 ACTIVATION PEPTIDE.
CC CC      FT PROPEP 16 59
CC CC      FT CHAIN 60 387 PEPsin II-4.
CC CC      FT MOD_RES 129 129 PHOSPHORYLATION (POTENTIAL).
CC CC      FT ACT_SITE 93 93 BY SIMILARITY.
CC CC      FT ACT_SITE 276 276 BY SIMILARITY.
CC CC      FT DISULFID 106 111 BY SIMILARITY.
CC CC      FT DISULFID 267 271 BY SIMILARITY.
CC CC      FT DISULFID 310 343 BY SIMILARITY.
CC CC      FT SEQUENCE 387 AA; 42052 MW; 21ADD07782A89585 CRC64;
CC CC
CC CC      Query Match 11.2%; Score 298; DB 1; Length 387;
CC CC      Best Local Similarity 26.1%; Pred. No. 1.8e-16;
CC CC      Matches 97; Conservative 66; Mismatches 122; Indels 86; Gaps 14;
CC CC
CC CC      75 YVEMTVGSPQPTNLIVDTGSSNFVGAAPHF-----LHRYQRLSSYRDLRKG 127
CC CC      75 YFTGISICTPQDFTVIFDTGSSNLW---PSTYCSLALCHKRFPEDSSYTGTS 131
CC CC
CC CC      128 VYVPTQCKWEGELGTLVSIPIHGNPNTVRANIAAITESDKFF-----INGSNWE 177
CC CC      132 LSITYGTGSMTGILGYDITV-----KVGSIEDTNIQIFGLSKTEPLTFAPED 179
CC CC
CC CC      178 GILGLGAYAEIARPDLSLEPFDFSLVKQTHV-PNLFSLHLCGAGFPLNQSEVLASVGGSMI 236
CC CC      180 GILGLAYPSISSDAT--PVFDNMWNEGLVSQDLFSVYLSDD-----EKGSLVM 227
CC CC
CC CC      237 IGGIDHSLYTGSLWYTPIRREWYEVIIIVRVEINGDLKM--DCKEYNDKSIDVDSGTTN 294
CC CC      228 FGGIDSSYTYTGLNWPVSYEGYQWITMDSVINGETIACADSC-----QAVDGTGSL 281
CC CC
CC CC      295 LRLPKKVFEEAAVKSKTAASTKEKFPDGFGLGEOLV-CWQAGTTPWNIFPVISLYLGEVNT 353
CC CC      282 LTGP-----TSAINIQSYIGASK-----NLLGENVISCAIDSLPDIVF----- 321

```

```
Qy 354 NQSFRTILPQQYLRPVEDVATSDCCYKFAISOSTGT-----VNGAVIMEGFYVVD 407
Db 322 -----TINGIQYPLPASAYILKEDDCTSGLEGMMVDTYTGTGELWGLDGFIRQYFTVD 375
Qy 408 RARKRIGFAVS 418
Db 376 RANQLGLAAA 386

RESULT 15
CATD_RAT CATD_RAT STANDARD; PRT; 407 AA.
AC P24268;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
CN CTSD.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Pituitary;
MEDLINE=91057150; PubMed=2243802;
RA Birch N.P., Loh Y.P.;
RT "Cloning, sequence and expression of rat cathepsin D.";
RL Nucleic Acids Res. 18:6445-6445(1990).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 65-74; 118-127 AND 165-174.
RC TISSUE=Liver;
MEDLINE=91354249; PubMed=1883350;
RA Fujita H., Tanaka Y., Noguchi Y., Kono A., Himeno M., Kato K.;
RT "Isolation and sequencing of a cDNA clone encoding rat liver
RT lysosomal cathepsin D and the structure of three forms of mature
RT enzymes.";
RL Biochem. Biophys. Res. Commun. 179:190-196(1991).
RN [3]
RP SEQUENCE OF 134-170.
RX MEDLINE=89034127; PubMed=3182800;
RA Yonezawa S., Takahashi T., Wang X., Wong R.N.S., Hartsuck J.A.,
RA Tang J.;
RT "Structures at the proteolytic processing region of cathepsin D.";
RJ J. Biol. Chem. 263:16504-16511(1988).
CC -!- FUNCTION: Acid protease active in intracellular protein breakdown.
CC -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC that of pepsin A. Does not cleave the 4-Gln-[His-5 bond in B
CC chain of insulin.
CC -!- SUBUNIT: OCCURS AS A MIXTURE OF BOTH A SINGLE CHAIN FORM AND TWO
CC TYPES OF TWO CHAIN (LIGHT AND HEAVY) FORMS.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54467; CAA38349.1; .
CC PIR; S13111; KHRTD.
CC PIR; JQ1177; JQ1177.
CC HSP; P07339; ILYB.
CC MEROPS; A01.009; -.
CC InterPro; IPR001461; AspproteaseA1.
CC InterPro; IPR001969; Aspprotease_site.
CC Pfam; PF00026; asp; 1.
CC PRINTS; PR00792; PEPSIN.
CC PROSITE; PS00141; ASP_PROTEASE; 2.
```

```
KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Signal;
Lysozyme.
FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 64 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 65 407 CATHEPSIN D.
FT CHAIN 65 164 CATHEPSIN D 12 KDA LIGHT CHAIN.
FT CHAIN 165 407 CATHEPSIN D 30 KDA HEAVY CHAIN.
FT CHAIN 65 117 CATHEPSIN D 9 KDA LIGHT CHAIN.
FT CHAIN 118 407 CATHEPSIN D 34 KDA HEAVY CHAIN.
FT ACT_SITE 97 97 BY SIMILARITY.
FT ACT_SITE 290 290 BY SIMILARITY.
FT DISULFID 91 160 BY SIMILARITY.
FT DISULFID 110 117 BY SIMILARITY.
FT DISULFID 281 285 BY SIMILARITY.
FT DISULFID 324 361 BY SIMILARITY.
FT CARBOHYD 134 134 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 15 15 D -> A (IN REF. 2).
FT CONFLICT 163 163 D -> T (IN REF. 3).
FT CONFLICT 205 205 K -> N (IN REF. 2).
FT CONFLICT 262 262 K -> N (IN REF. 2).
SQ SEQUENCE 407 AA; 44680 MW; C423AD4104D95F84 CRC64;

Query Match 11.1%; Score 297; DB 1; Length 407;
Best Local Similarity 26.1%; Pred. No. 2.3e-16;
Matches 118; Conservative 76; Mismatches 170; Indels 88; Gaps 20;

Qy 6 PWLLWMGAGVLPAGHTGHIRPLR-----SGLGGA--PLGURLPRETDEEPEP 54
Db 4 PGVLLLI-LGLDASSAL-IRPLRKFTSIRMTVEVGSGVEDLILKPTIKYMOSSP 61

Qy 55 GRGSEFVEMVDNLGRKSGQGYVEMTVGSPQTLNLILVDTGSSNFAVGAAPHPL----- 109
Db 62 RTKEPVSELLKNVLDQ---YYEIGIGTPQCFTVVFDTGSSNLWVPSIHCCKLLDIACW 118

Qy 110 -HRYQRLSSYTRDLRKGVVYPTQCKWEGELGTDLVSPHGPNTVTRVANIAITESDK 168
Db 119 VHHKYNDSKSTYVKNGTSEFDIHYGSGSLSGYLSQDTSVSP-----CKSDLGIGIKVERQ 172

Qy 169 FF-----INGSNWEGILGLAYAEIARPDSDLEPFDLSLVKQTHV-PNLFSLHLCG 217
Db 173 IFGEATKQPGVFIAAKFDGILGNGYPPIS--VNKVLVPFDNLKMKQLVKNIFSFYL-- 228

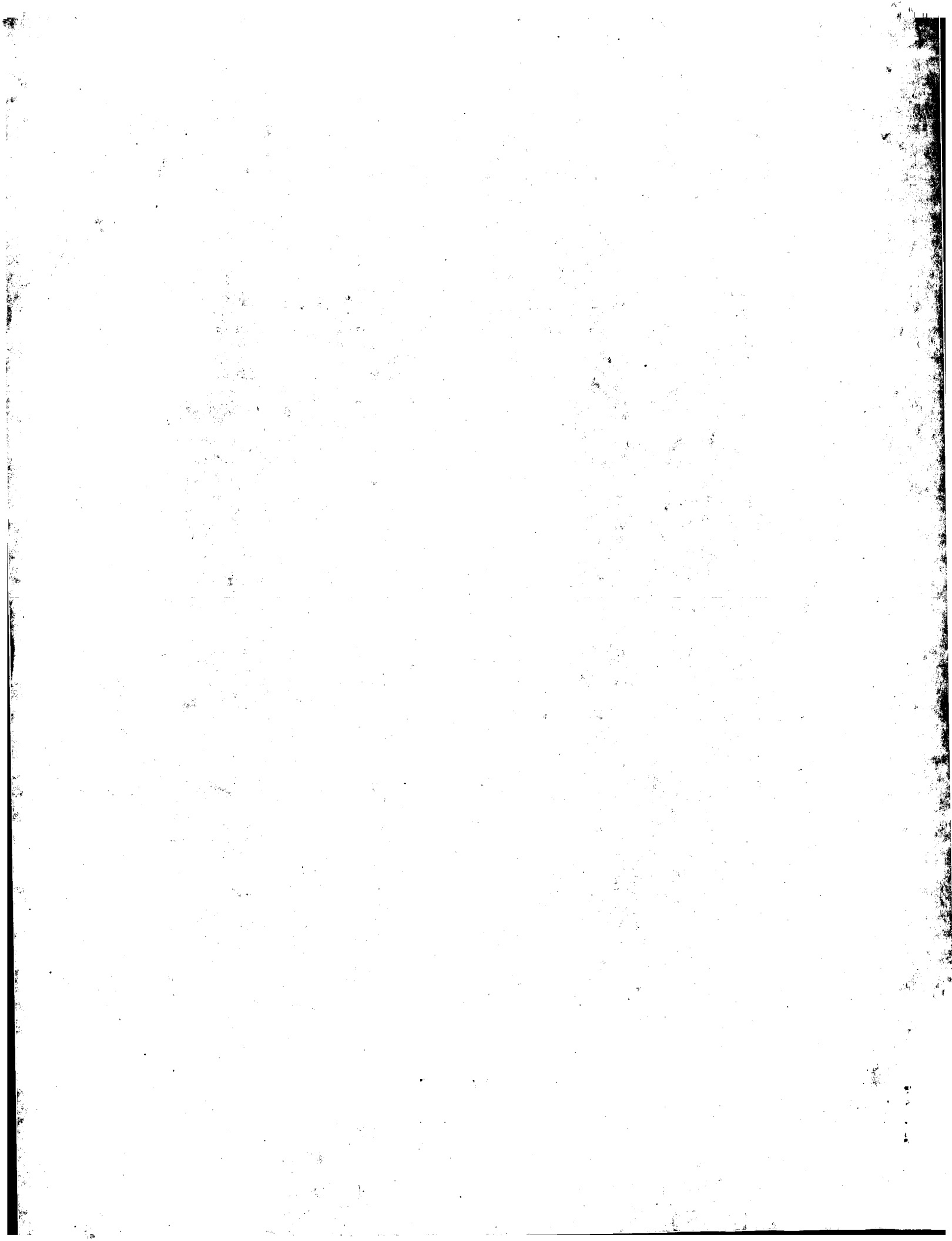
Qy 218 AGFPLAQSEVLASVGGSMIIIGIDHSLYTGLSWYTPIRREWYVEIIVRVEINGQDLKMD 277
Db 229 -----NRDPTGQPGGELMLGGTDSRYVHCELSVNLVTRRAYQVHMDQLEV-GSEULT- 280

Qy 278 CKEYNDKSTVDSGTTNLRPKVFEAAVKSIAASSTKEKFPDGFGLGQLEV-CWQAGTT 336
Db 281 CK--GCEAIVDTCTSLLVGPVDEVKELQKAIGAVPLIQ-----GEYNIPEKRVSS- 329

Qy 337 PMNIFPVISLYLMGEVNTQSFRTILPQOYLRPVEDVATSDCCYKFAIS-----Q 387
Db 330 ----LPIITFKLGQ-----NYELIPEKYLKVSQAGKT-----ICLSGFMGMDIPPP 373

Qy 388 SSTGTVMCAVIMEGFYVVDRAKRIGFAVSA 419
Db 374 SGPLWLIGDVFICGYTVTFDREYNRVGFAKAA 405

Search completed: June 5, 2003, 16:12:44
Job time : 24 secs
```



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:11:35 ; Search time 43 seconds
(without alignments)
1120.078 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMMGAGVLPAAH.....CLRCLRQHQHDDFADDISLLK 501
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Aligned: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2656	99.7	501	2 A59090	aspartic proteinase
2	324	12.2	384	2 JC7574	pepsinogen A - Afr
3	313.5	11.8	385	2 JC7575	pepsinogen A - bul
4	309	11.6	387	2 B38302	pepsin (EC 3.4.23)
5	307.5	11.5	383	2 JC7573	pepsinogen C - Afr
6	307.5	11.5	388	2 SI9682	pepsin A (EC 3.4.2)
7	305	11.4	384	2 A39314	gastricsin (EC 3.4
8	304	11.4	382	1 PECH	pepsin A (EC 3.4.2
9	301.5	11.3	383	2 A41443	pepsin (EC 3.4.23
10	301.5	11.3	396	2 A3401	cathepsin E (EC 3.
11	300.5	11.3	412	1 KHUD	cathepsin D (EC 3.
12	299	11.2	387	2 C38302	pepsin (EC 3.4.23
13	298	11.2	387	2 D38302	pepsin (EC 3.4.23
14	297	11.1	407	1 KHRTD	cathepsin D (EC 3.
15	295	11.1	391	2 A43356	cathepsin E (EC 3.
16	294.5	11.1	388	1 SI9684	pepsin A (EC 3.4.2
17	290.5	10.9	380	2 I47176	chymosin (EC 3.4.2
18	289	10.8	387	2 E38302	pepsin (EC 3.4.23
19	288.5	10.8	388	1 PEHU	pepsin A (EC 3.4.2
20	288.5	10.8	388	1 PEMQAR	pepsin A (EC 3.4.2
21	288.5	10.8	388	2 A30142	pepsin A (EC 3.4.2
22	288.5	10.8	398	2 S66455	cathepsin E (EC 3.
23	288	10.8	389	2 JE0371	pepsin C (EC 3.4.2
24	287.5	10.8	444	2 T24204	hypothetical prote
25	287	10.8	410	1 KHMSD	cathepsin D (EC 3.
26	286.5	10.8	388	1 PEMQAJ	pepsin A (EC 3.4.2
27	286	10.7	398	2 I51185	cathepsin D (EC 3.
28	285.5	10.7	388	2 B30142	pepsin A (EC 3.4.2
29	284.5	10.7	381	1 CMShB	chymosin (EC 3.4.2

ALIGNMENTS

RESULT 1

A59090
aspartic proteinase (EC 3.4.23.-) BACE precursor - human
N:Alternate names: beta-secretase; beta-site APP cleaving enzyme
C:Species: Homo sapiens (man)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000
C:Accession: A59090
R:Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.; Tello
M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Ro
Science 286, 735-741, 1999
A:Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the tran
A:Reference number: A59090; MUID:2002972; PMID:10531052
A:Note: submitted to GenBank, September 1999
A:Accession: A59090
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-501 <VAS>
A:Cross-references: GB:AF190725; NID:G6118538; PIDN:AAF04142.1; PID:G6118539
C:Genetics:
A:Gene: BACE
C:Superfamily: beta-secretase
C:Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein; hydrolase;
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-45/Domain: propeptide #status predicted <PRO>
F:46-501/Product: acid proteinase BACE #status predicted <MAT>
F:461-477/Domain: transmembrane #status predicted <TRN>
F:93-289/Active site: Asp #status predicted
F:153,172,223,354/Binding site: carboxylate (Asn) (covalent) #status predicted
F:330-380/Disulfide bonds: #status predicted

Query Match	99.7%	Score 2656;	DB 2;	Length 501;
Best Local Similarity	99.8%	Pred. No. 7e-206;		
Matches 500;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MAQALPWLMMGAGVLPAAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEPCRGSGF	60	
Db	1	MAQALPWLMMGAGVLPAAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEPCRGSGF	60	
Qy	61	VEVDNLRGKSGQGYVEMTVGSPQTLNVLDTGSSNFVGAAPHPFLHRYQRQLSST	120	
Db	61	VEVDNLRGKSGQGYVEMTVGSPQTLNVLDTGSSNFVGAAPHPFLHRYQRQLSST	120	
Qy	121	YRDLRKGYVPYTGKWEGLGTDLVSTPHGPNVTVRANTAAITDSKFFFTNGWEGIL	180	
Db	121	YRDLRKGYVPYTGKWEGLGTDLVSTPHGPNVTVRANTAAITDSKFFFTNGWEGIL	180	
Qy	181	GLAYAEIARPDSDLPEPFDSLVKQTHVNPFLSLHLCGAGFPPLNQSEVLASVGGSMIIGGI	240	
Db	181	GLAYAEIARPDSDLPEPFDSLVKQTHVNPFLSLHLCGAGFPPLNQSEVLASVGGSMIIGGI	240	
Qy	241	DHSLYTGSLWYTPTRREWYVEVITVRVEINGQDLKMDCKEYNDKSVDSGTTNLRLPKK	300	
Db	241	DHSLYTGSLWYTPTRREWYVEVITVRVEINGQDLKMDCKEYNDKSVDSGTTNLRLPKK	300	

pepsin A (EC 3.4.2
pepsinogen A - com
chymosin (EC 3.4.2
cathepsin E (EC 3.
aspartic proteinase
gastricsin (EC 3.4
prochymosin - comm
pepsin (EC 3.4.23
cathepsin D (EC 3.
aspartic proteinase
pepsinogen C - com
gastricsin (EC 3.4
saccharopepsin (EC
gastricsin (EC 3.4
pepsin A (EC 3.4.2
candidapepsin (EC

```

316 IVF-----TINGVQYPLSPSAVYRQNOGGSSGFQAMNLTNSGDLW 357
393 VMGAVIMEGFYVDFDRARKRIGFA 416
358 ILGDVFIROYFTVDFRANNVYVIA 381

RESULT 3
JC7575
pepsinogen A - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C:Accession: JC7575
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: purification and characterization of Xenopus pepsinog
A:Reference number: JC7573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: JC7575
A:Molecule type: mRNA
A:Residues: 1-385 <IKU>
A:Cross-references: DDBJ:AB045376
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-lik
C:Genetics:
C:Gene: Pga
C:Superfamily: pepsin
C:Keywords: stomach; zymogen

Query Match 11.8%; Score 313.5; DB 2; Length 385;
Best Local Similarity 26.6%; Pred. No. 2.2e-17;
Matches 117; Conservative 74; Mismatches 158; Indels 91; Gaps 20;

QY 8 LLLWMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEGRGSEV-----61
DB 3 ILLFLGLVLAECGV---VKVSLRK---GESLRARLNR-----LGLLDYLLKHHYN 48
QY 62 ---EMVDNLKRGSGO-----GYVEMTVGSPPTNLILVDTSNFAVG-----AAPH 106
DB 49 PATKYFPFSAQAAGEPLQNMVMDIEYGTISIGTPPQSFVIFDTGSSNLWVPSVYCSSPA 108
QY 107 PFLHRYQRLSSTYRDLRGVYVPTQGWEGELGTLVSLPHGPNVIVRANIAITES 166
DB 109 CTNHMFENPOOSSTFQATNTPVSIQYGTGSMGFLGYDTLVQV---GNIQITNQIFLSQS 165
QY 167 DK-FFINGSNWEGILGLAYAEIARPDLSLEPFDFSLVKOTHVP-NLFSLHLCGAGPLNQ 224
DB 166 EPGSFLYSPFDGLGLAFPSLA--SSQATPVFDNMWQGLIPQDLFSVYLSQG---Q 219
QY 225 SEVLASVGGSMIGGIDHSLYTGLSWYTPIRREWYVEIIVRVEINGQDLKM--DCKEYN 282
DB 220 S-----GSFVFGVDTSYTGTLNWNVPLTAETWQITVDSISIGQVIACSGSC---269
QY 283 YDKSIVDSGTTNLRPLPKVFEAAVKSIAKASSTKFPDGFMLGEOLVCMQAGTTPWNIFP 342
DB 270 --SAIVDTGTSLLAGP-----STPIANIQYIGANQDSNGQYV---INCNNISNMPVVF- 319
QY 343 VISLYLMEVTNQSFRTILPQOYLRPVED--VATSDQDC---YKFAISQSSTGT--VMGA 396
DB 320 -----TINGVQYPLPASAVYRQSQSCTSGFQAMNLTSSGDLWILGD 362

397 VIMEGFYVDFDRARKRIGFA 416
363 VFIREYVYVDFRANNVYVIA 382

RESULT 4
B38302
pepsin (EC 3.4.23.-) II-1 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Jun-1991 #sequence_revision 20-Sep-1991 #text_change 23-Feb-1997
C:Accession: B38302
R:Kageyama, T.; Tanabe, K.; Koiwai, O.
J. Biol. Chem. 265, 17031-17038, 1990

```


157 157 AATAAATESDK-FFINGSNWEGILGLAYAEIARPDSDLEPPFDSLVKQTHV-PNLFSJH 214
159 159 TNQIFGLSETEPEFFLYAFDGLGLAYPSIS--SSGATPVFDNIWNRQLVNSQDLFSVY 216
215 215 LCGAGFLPNQSEVLASVGGSMITGGIDHSILYTSGLWYTPIRREWEYEVIIIVRVEINGQDL 274
217 217 LSAD-----DOS-----GSVIFGGIDSSYITGSLANWVPVSEGYWQISVDSITMGKTI 266
275 275 --KMDCKEYNDKSIDVSGTTLNLRPKKVEAAVKSIIKAASSTKPKPDGFWLGEQLV-CW 331
267 267 ACAGC-----QAIVDTGSLTGTPTSPANTQSDIGASENSD-----GEMVVS 312
332 332 QAGTTPWNPFFVISLYLMEVNTSFRITILPOOY-LRPVEDVATSDDCYK-----FAI 385
313 313 AISSLPDIVF-----TINGVOYPLPPSAYILQSGSCTSGFQGMVDP 354
386 386 SOSSTGTVMGAVIMEGFVYVFDRAKRIGFA 416
385 385 TESGELWILGDVFIROYFTVFDRAANNQVGLA 385

RESULT 7
A39314
gastricsin (EC 3.4.23.3) precursor - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 22-Jun-1999
C:Accession: A39314
R:Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kageya
J. Biol. Chem. 266, 22436-22443, 1991
A:Title: Purification, characterization, and amino acid sequences of pepsinogens and pep
A:Reference number: A39314; MUID:92042186; PMID:1939266
A:Accession: A39314
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <XAK>
A:Cross-references: GB:M73750; NID:g213687; PIDN:AAA49530.1; PID:g213688
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 11.4%; Score 305; DB 2; Length 384;
Best Local Similarity 24.5%; Pred. No. 1.le-16;
Matches 105; Conservative 66; Mismatches 146; Indels 112; Gaps 17;

QY 24 HGIRLRLSLGAGLGLRLPRETDEPEPEGR--GSFVEMVDNLRGKSGOGYVEMTV 81
DB 35 HGIRAPV-----VDPATKYNNFATAFEPLANVMDMSYVGEISI 73

QY 82 GSPPTNLILVDVTGSSNFVAGAAPHPFL-----HRYQROLSTYRDLRGVVPYQ 134
DB 74 GTPPQNFVLFDTGSSNLWV--PSTYCQSQACTNHPQNPQSSSYSSNQOQFSLQYGT 130

QY 135 GKVEGELGTLVSIHPGPNVTRANTA-----AITESDKFFINGSNWEGILGLAYAE 186
DB 131 GSUTGILGIDYTVQI-----QNAISQOBEGLSVTEGTFNFIY-AQFDGILGLAYS 180

QY 187 IARPDSDLEPPFDSLVKQTHV-PNLFSHLCGAGFLPNQSEVLASVGGSMITGGIDHSILY 245
DB 181 IA--EGGATVMQMIQNLINQPLFAFLVSGOONSQ-----GGEVAFEGVDQNY 230

QY 246 TGSWYTPIRREWEYEVIIIVRVEINGQD---LMDCKEYNDKSIDVSGTTLNLRPKKVF 302
DB 231 SGQIYTPVTSYQYTGIOGFGVNGOATGWCSCG-----QGIVDITGSLTAPQSYF 284

QY 303 EAAVKSKIAASSTKPKPDGFWLGEQLV-CWQAGTTPWNPFFVI-----SLYLMEVTV 353
DB 285 SSLMQSIGAQDQD-----GOYAVSCSNISQSLPTISTISGVFPPLPPSAVILQONS 336

QY 354 NQ-----SFRITILPOOYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGFVYVFDRA 410
DB 337 GYCTIGTMTPTVLPSONQPL-----WILGDVFLRQYISYVYDLGN 375

QY 411 KRIGFAVSA 419

376 376 NOVGFAAAA 384

RESULT 8
PECH
pepsin A (EC 3.4.23.1) precursor - chicken
N:Alternate names: pepsinogen A
C:Species: Gallus gallus (chicken)
C:Date: 18-Apr-1984 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C:Accession: JE0370; A00984
R:Sakamoto, N.; Saiga, H.; Yasugi, S.
Biochem. Biophys. Res. Commun. 250, 420-424, 1998
A:Title: Analysis of Temporal expression pattern and cis-regulatory sequences of chic
A:Reference number: JE0370; MUID:98440813; PMID:9753645
A:Accession: JE0370
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-382 <SAK>
A:Cross-references: GB:AB025281; NID:g4589837; PIDN:BAA76891.1; PID:g4589838
R:Baudys, M.; Kostka, V.
Eur. J. Biochem. 136, 89-99, 1983
A:Title: Covalent structure of chicken pepsinogen.
A:Reference number: A00984; MUID:84004412; PMID:6617663
A:Accession: A00984
A:Molecule type: protein
A:Residues: 16-87, 'S', 89-382 <BAU>
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; gastric juice; glycoprotein; hydrolase; protein dige
F:16-57/Domain: activation peptide status experimental <APT>
F:58-382/Product: pepsin A #status predicted <MAT>
F:92,275/Active site: Asp #status predicted
F:105-110,266-270,305-338/Disulfide bonds: #status experimental
F:128/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 11.4%; Score 304; DB 1; Length 382;
Best Local Similarity 24.0%; Pred. No. 1.3e-16;
Matches 88; Conservative 69; Mismatches 125; Indels 84; Gaps 13;

QY 75 YVEMTVGSPPTNLILVDVTGSSNFVAGAAPHPFL-----HRYQROLSTYRDLRGK 127
DB 74 YGTFISICTPQDFTVIFDTGSSNLWV--PSTYCKSSACSNHCRDPSKSYVSTNET 130

QY 128 VVYPYTOGKVEGELGTLVSIHPGPNVTRANTA-----HRYQROLSTYRDLRGK 186
DB 131 VYTAGTSGMSGILGYDTAV---SSIDVQNLQIFGLSETEPGSFYFFYCNFDGLGLAFPS 187

QY 187 IARPDSDLEPPFDSLVKQTHV-PNLFSHLCGAGFLPNQSEVLASVGGSMITGGIDHSILY 245
DB 188 IS--SSGATPVFDNMMSQHLVAQDLFSVLSKDG-----ETGSFVLFQGDIPNYT 235

QY 246 TGSWYTPIRREWEYEVIIIVRVEINGQDLK-MDCKEYNDKSIDVSGTTLNLRPKKVF 303
DB 236 TKGIYVWPLSAETWQITMDRVTVGNKYVACFFTC-----QAIVDTGSLVMPQAYN 289

QY 304 AAVSKIAASSTK-----KFPDGFVLGEQLVQWQAGTTPWNPFFVISLYLMEVNTQS 356
DB 290 RIIRDGLVSSDGEISCDISKLPD-----VTFHINGHA----- 322

QY 357 FRITILPOOYLRPVEDVATSDDCYKFAISOSSTGT-----VMGAVIMEGFVYVFDRA 410
DB 323 -----FTLPASAVVLNEDGSCMLGFENMGTPTELGEQWILGDVFIREYVIFDRA 373

QY 411 KRIGFA 416
DB 374 NKVGLS 379

RESULT 9
A41443
pepsin (EC 3.4.23.-) precursor, embryonic - chicken
C:Species: Gallus gallus (chicken)
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 21-Jul-2000

R: Hill, J.; Montgomery, D.S.; Kay, J.
FEBS Lett. 326, 101-104, 1993
A: Title: Human cathepsin E produced in E. coli.
A: Reference number: S34467; MUID: 93314762; PMID: 8325357
A: Accession: S34467
A: Status: preliminary
A: Molecule type: protein
A: Residues: 57-60, 62-81 <HIL>
R: Athauda, S.B.P.; Matsuzaki, O.; Kageyama, T.; Takahashi, K.
Biochem. Biophys. Res. Commun. 168, 878-885, 1990
A: Title: Structural evidence for two isozymic forms and the carbohydrate attachment s
A: Reference number: A34643; MUID: 90241267; PMID: 2334440
A: Accession: A34643
A: Status: preliminary
A: Molecule type: protein
A: Residues: 54-58, 'XXX', 62-64, 'M', 66-89, 'X', 91-95 <ATH>
A: Accession: B34643
A: Status: preliminary
A: Molecule type: protein
A: Residues: 54-59, 'X', 61-68 <AT2>
C: Genetics:
A: Gene: GDB: CTSE
A: Cross-references: GDB: 119821; OMIM: 116890
A: Map position: Iq31-Iq31
C: Superfamily: pepsin
C: Keywords: aspartic proteinase; blocked amino end; hydrolase; zymogen
F: 1-17/Domain: signal sequence #status predicted <SIG>
F: 18-53/Domain: activation peptide #status predicted <PRO>
F: 54-396/Product: cathepsin E #status predicted <MAT>
F: 18/Modified site: blocked amino end (Gln) (in mature form) (probably pyrrolidone ca
F: 96, 281/Active site: Asp #status predicted

Query Match 11.3%; Score 301.5; DB 2; Length 396;
Best Local Similarity 25.88; Pred. No. 2.1e-16;
Matches 100; Conservative 68; Mismatches 144; Indels 75; Gaps 16;

Qy 48 DEPEEPGRGRSGFVEMVDNLGRKSGQGYVEMTVGSPQPTLNILVDTSNFAVGA---- 103
Db 63 DQSAKEP-----LINYLD-----MEYFGTISIGSPQNFVFDTGSSNLWVPSVYCT 110
Qy 104 APHFLHRYQRLSTYRDLRGVVPYVYQGWEGELGTDLVSIHPGPNVTVRANIAAI 163
Db 111 SPACKTHSRFPQSSSTPYSPQGSFSIQYGTGSLGIIGADQVSV-EGLTVMGQOFGESV 169
Qy 164 TESDKPFINSNEGILGLAYAEIARDDDSLEPFDSLVKQTHVP-NLFSHLHLCGAGFPL 222
Db 170 TEPQTFVD-AEFDGLIGLYPSLA--VGGVTPVFDNMMAQNLVDLPMSFVYM----- 219
Qy 223 NQSEVLASVGSMSIIGDIHSLTSGSLWYTPIRREWYEVIIVRVEINGODLKMDCKEYN 282
Db 220 -SSNPEGAGSELIFGQYDHSFSGSLNWVPVTKQAYWQIALDNIQVGG--TVMFCSE-- 274
Qy 283 YDKSIVDSGTTNLRLPKKVFEAAVKSTKAASSTTEKPPDGFGLGQLVCWQAGTTPWNIFP 342
Db 275 GCGAIVDTGTLTSPGSDKIKQLQNAIGAAP-----VDGEYAVE-----CANLNVMP 321
Qy 343 VISLYLMGEVNTNOSFRITILPQOYLRPEVDVATSQDDCYKFAISQSTG----- 391
Db 322 DVFTTNG-----VPYTLSTPAY--TLLDFVDMGMQFC-----SSGFGQLDIHPAG 365
Qy 392 --TVMGAVIMEGFVYVFDRAKRRTGFA 416
Db 366 PLWILGDVFIRQFYSVFDRCNNRVGLA 392

RESULT 11.
KHHUD
N: Alternate names: preprocathepsin D
C: Species: Homo sapiens (man)
C: Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 15-Sep-2000
C: Accession: A25771; S30749; PC2066; I59236; I57716
C: Faust, P.L.; Kornfeld, S.; Chirowin, J.M.

Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985
A:Title: Cloning and sequence analysis of cDNA for human cathepsin D.
A:Reference number: A25771; MUID:85270436; PMID:3927292
A:Accession: A25771
A:Molecule type: mRNA
A:Residues: 1-412 <FAD>
A:Cross-references: EMBL:M11233; NID:g181179; PIDN:AAB59529.1; PID:g181180
R:Westley, B.R.; May, F.E.B.
Nucleic Acids Res. 15, 3773-3786, 1987
A:Title: Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human breast
A:Reference number: S30749; MUID:87231068; PMID:3588310
A:Accession: S30749
A:Molecule type: mRNA
A:Residues: 1-412 <NES>
A:Cross-references: EMBL:X05344; NID:g29677; PIDN:CAA28955.1; PID:g29678
R:May, F.E.B.; Smith, D.J.; Westley, B.R.
Gene 134, 277-282, 1993
A:Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulated
A:Reference number: PC2066; MUID:94085791; PMID:8262386
A:Accession: PC2066
A:Molecule type: DNA
A:Residues: 1-23 <MAY>
A:Cross-references: GB:112980; NID:g291930; PIDN:AAA16314.1; PID:g455429
A:Experimental source: MCF-7 cell
R:Cavaillès, V.; Augereau, P.; Rochefort, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993
A:Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only
A:Reference number: I59236; MUID:93126342; PMID:8419924
A:Accession: I59236
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <CAV1>
A:Cross-references: GB:S5257; NID:g263124; PIDN:AAD13868.1; PID:g4261568
R:Augereau, P.; Miralles, F.; Gaudelot, C.; Parker, M.; Rochefort, H.
Mol. Endocrinol. 8, 693-703, 1994
A:Title: Characterization of the proximal estrogen-responsive element of human cathepsin
A:Reference number: I57716; MUID:95021301; PMID:7935485
A:Accession: I57716
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <CAV2>
A:Cross-references: GB:S74689; NID:g786350; PIDN:AAD14156.1; PID:g4261856
R:Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A:Reference number: A51839; PDB:1LYA
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161;170-241
R:Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A:Reference number: A51840; PDB:1LYB
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, with inhibitor residues 65-
R:Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Hosur, M.V.; Sowder II, R.C.; Cachau, R.E.; Col
Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993
A:Title: Crystal structures of native and inhibited forms of human cathepsin D: implicat
A:Reference number: A48229; MUID:93342076; PMID:8393577
A:Contents: annotation; X-ray crystallography, 2.5 angstroms
C:Comment: Cathepsin D is a ubiquitous lysosomal proteinase.
C:Comment: In addition to the propeptide, residues 163-168 and 411-412 are proteolytical
C:Comment: The carbohydrate bound to 134-Asn contains a mannose-6-phosphate that is bound
C:Genetics:
A:Gene: GDB:CTSD
A:Cross-references: GDB:120512; OMIM:116840
A:Map position: l1p15.5-l1p15.5
C:Function:
A:Description: limited specificity endopeptidase
A:Pathway: intracellular protein degradation
C:Superfamily: peptin
A:Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-64/Domain: propeptide #status predicted <PRO>
F:65-162/169-410/product: cathepsin D #status experimental
F:267-329-356/Region: phosphotransferase recognition
F:91-160-110-117-286-290-329-366/Disulfide bonds: #status experimental
F:97, 295/Active site: Asp #status experimental

F:134,263/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 11.2%; Score 300.5; DB 1; Length 412;
Best Local Similarity 26.9%; Pred. No. 2.7e-16;
Matches 123; Conservative 68; Mismatches 170; Indels 97; Gaps 21;
QY 5 LPWLLWMGAGVLPAGHTGIRLPLR-----SGLGGAPLGL-----RLP 44
Db 7 LPALALCAA---PASAL---VRPLHKFTSIRRTMSEVSGSVEDLIARQPVSKYQAVP 60
QY 45 RETDEPEPEPRGRGSFVMDNLRGSGQGYVEMVGPPTQTLNLTVDTSNFAVGA 104
Db 61 AVTE-----GPIPEVLKNYMDAQ---YGEIGIGTTPPOCTVVFEDTGSNLWVPSI 108
QY 105 PHPEL-----HRYORQLSSTYRDLRKGVVYPTQCKWEGELGTDLVSTP----- 149
Db 109 HCKLLDTACWIHKYNSDKSSTVYKNGTSFDIHYGSGLSGLYSQDTSVSPQCSASSASA 168
QY 150 HGPVNTVRANIAAITESSDKFFINGSNWEGILGLAYABIAIPDDSLRFFDLSLVKQTHV-P 208
Db 169 LGGVKVERQVFEATKQGITFAAKFDGLIGMAYPRIS--VNNVLPVFDNMQOKLVDQ 226
QY 209 NLFSLHLCGAGFPLNQSEVLASVGSMTIGGIDHSLVTGSLWYTPIRREWYEVIIVRVE 268
Db 227 NIFSFL-----SRDPAQPGGELMGLGTDSSKYKGSLSYLNTRKAYWQVHLDQVE 278
QY 269 I-NGQDLKMDCKEYNDKSIIVDSGTTNLRPLPKKVFEEAAVKSIAASSTKPPGFWLGEQ 327
Db 279 VASGLTL---CKE--GCEAIVDTGTSMLMVGPDVRELOKAIGAVPLIQ-----GEY 325
QY 328 LV-CWQAGTTWNIFPVLSLYLMGEVYNQSFRTILPQOYLRPVEDVATSDQDCYKFAIS 386
Db 326 MIPCEKVT-----LPAITLKLGG---KGYKLS--PEDTYLVKSQAGKTL---CLSGFMG 372
QY 387 Q-----STGTVMGAVIMEGFVYVVDRAKRKRIGFAVSA 419
Db 373 MDIPPPGFLMLGDFVFGYTYTVFDRDNNRVGFAEA 410
RESULT 12
C38302
pepsin (EC 3.4.23.-) II-2/3 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 23-Feb-1997
C:Accession: C38302
R:Kageyama, T.; Tanabe, K.; Koiwai, O.
J. Biol. Chem. 265, 17031-17038, 1990
A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nu
A:Reference number: A38302; MUID:91009127; PMID:2129536
A:Accession: C38302
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-387 <KAG>
A:Cross-references: GB:J05638
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion
Query Match 11.2%; Score 299; DB 2; Length 387;
Best Local Similarity 26.9%; Pred. No. 3.3e-16;
Matches 97; Conservative 64; Mismatches 134; Indels 66; Gaps 13;
QY 75 YVEMTVGSPPTQTLNLTVDTSNFAVGAAPHPF-----LHRYORQLSSTYRDLRKG 127
Db 75 YFGTISIGTPPODFTVFEDTGSNLW---PSTYCSLLACALHKRFNPEDSDTQGTST 131
QY 128 VYVYTGCKWEGELGTDLYSIPHGPNVTVRANIAAITESSDKFFINGSNWEGILGLAYABE 187
Db 132 LSTYTGSGMTGLGYDTRVKVGSIEDTNQIFGLSKTEPSLTFE--APDGILGLAYPSI 189
QY 188 ARPDLSLEPPFDLSLVKQTHV--PNLSLHLCGAGFPLNQSEVLASVGSMTIGGIDHSLYT 246
Db 190 SSSDAT--PVFDNNWNEGLVSQDLFSVYLSSDD-----EKGSLVMFGGIDSSYYT 237

C:Accession: S1311; C31918; JQ1177; PQ0222
R:Birch, N.P.; Loh, Y.P.
Nucleic Acids Res. 18, 6445-6446, 1990
A:Title: Cloning, sequence and expression of rat cathepsin D.
A:Reference number: S1311; MUID:91057150; PMID:2243802
A:Accession: S13111
A:Molecule type: mRNA
A:Residues: 1-407 <BIR>
A:Cross-references: EMBL:X54467; NID:g55881; PIDN:CAA38349.1; PID:g55882
R:Yonezawa, S.; Takahashi, T.; Wang, X.; Wong, R.N.S.; Hartsuck, J.A.; Tang, J.
J. Biol. Chem. 263, 16504-16511, 1988
A:Title: Structures at the proteolytic processing region of cathepsin D.
A:Reference number: A92681; MUID:89034127; PMID:3182800
A:Accession: C31918
A:Molecule type: protein
A:Residues: 134-162, 'T', 164-170 <VON>
R:Fujita, H.; Tanaka, Y.; Noguchi, Y.; Kono, A.; Himeno, M.; Kato, K.
Biochem. Biophys. Res. Commun. 179, 190-196, 1991
A:Title: Isolation and sequencing of a cDNA clone encoding rat liver lysosomal cath
A:Reference number: JQ1177; MUID:91354249; PMID:1883350
A:Accession: JQ1177
A:Molecule type: mRNA
A:Residues: 1-14, 'A', 16-204, 'N', 206-261, 'N', 263-407 <FUJ>
A:Accession: PQ0222
A:Molecule type: protein
A:Residues: 65-74; 118-127; 165-174 <FU2>
A:Experimental source: liver
C:Comment: Cathepsin D in rat liver lysosome occurs as a mixture of both a single cha
C:Function:
A:Description: limited specificity endopeptidase,
A:Pathway: intracellular protein degradation
C:Superfamily: pepsin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-64/Domain: propeptide #status predicted <PRO>
F:65-407/Product: cathepsin D, 43k single-chain form #status predicted <MAT>
F:65-164/Product: (or 65-165) cathepsin D 12K light chain #status predicted <MA2>
F:65-117/Product: cathepsin D 9K light chain #status predicted <MA4>
F:118-407/Product: cathepsin D 34K heavy chain #status predicted <MA5>
F:165-407/Product: (or 166-407) cathepsin D 30K heavy chain #status predicted <MA3>
F:91-160, 110-117, 281-285, 324-361/Disulfide bonds: #status predicted
F:97, 290/Active site: Asp #status predicted
F:134, 258/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.1%; Score 297; DB 1; Length 407;
Best Local Similarity 26.1%; Pred. No. 5.le-16;
Matches 118; Conservative 76; Mismatches 170; Indels 88; Gaps 20;

QY 6 PWLLLMGAGVLPAGHTQHRIPLR-----SGLGGA--PLGLRLPRETDPEEP 54
DB :|||:|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 4 PGVLLLI-LGLLDASSAL-IRIPLRKFTSIRRTMEVGSGVEDLILKGPITKYSMQSSP 61
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 55 GRGGSFVMVDNLRGKSGGYIVEMTVGSPQTLNLIVDTGSSNFVAAGRPHFL----- 109
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 62 RTKEPSVELLNKYDLAQ---YYGEIGITGPQCFTVFYDGTGSSNLNVSPSTHCKLLDIAC 118
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 110 -HRYQRQLSSYTRDLRKGVYPYPTQGWEGELGTDLSIPHGNVTVRANTAAITESDK 168
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 119 VHHKYNSDKSSHVVKNKTFTSDIHGSGSLSGYLSDQTVPV-----CKSDLGGIKVEKQ 172
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 169 FF-----INGSNWEGILGAYAEIARPDSDLPPFDLSLVKOTHY-PNLFSLHLGC 217
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 173 IFGEATKQPGVVFIAAKFDGILGMGYFFIS--VNKVLVPFDNLMMKKQLKFNIFSFL-- 228
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 218 AGFPLNQSEVLASVSGSMIIGGIDHSLYTGSLMYTPIRREWHYEVIIVRVENQDLKMD 277
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 229 -----NRDPTGQPGGELMLGGTDSRYHYHGELSYLNVRTRKAYQVHMDDLEV-GSELTL- 280
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 278 CREYNVDKSIVDSGTTNLRPKKVFEAAVKSKAAASTEKFPDPGFWLGEOLV-CWQAGTT 336
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 281 CK--GGCEALVDGTSLLVGPDVEKLOKAIGNVLIQ-----GETMIPCEKVS- 329
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 337 PWNIFPVISLYLMGEVNTNOSFRITILPOQYLRPEVDVATSQDDCYKFAIS-----Q 387
DB :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

QY      278 CKEYNYDKSIVDSGTTNLRPKKVFEEAAVKSIAASSTEFDPDFGLGQLV-CWQAGTT 336
        ||::|||::| | : | | : | | :: | : | :
Db      281 CK--GGCEAIVDTGTSLLVGPVDVKELKAIGAVPLIQ-----GEWMIPCEKVSS- 329

QY      337 PNMFIFVLSLYLMGEVTNSFRITILPOOYLRPVEDVATSQDDCYKFAIS-----Q 387

```

Fri Jun 6 08:31:00 2003

Db 330 ----LPIITFKLGQ-----NYELHPEKYLKVSQAGKT-----ICLSGFMGMDIPPP 373
QY 388 SSGTGVMGAVIMEGFYVVDRAKRGIFAVSA 419
Db 374 SGPLWILGDFVIGCYTYVFDREYNRVGFAKAA 405

RESULT 15
A43356
cathepsin E (EC 3.4.23.34) precursor - guinea pig
N:Alternate names: erythrocyte membrane aspartic proteinase; slow-moving proteinase
C:Species: Cavia porcellus (guinea pig)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: A43356
R:Kageyama, T.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Koiwai, O.; Tanji, M.
J. Biol. Chem. 267, 16450-16459, 1992
A:Title: Gastric procathepsin E and progastricsin from guinea pig. Purification, molecular
A:Reference number: A43356; MUID:92355614; PMID:1644829
A:Accession: A43356
Molecule type: mRNA
Residues: 1-391 <KAG>
A:Cross-references: GB:M88653; NID:g191294; PIDN:AAA37052.1; PID:g191295
A:Note: sequence extracted from NCBI backbone (NCBIN:110763, NCBI:P:110769)
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; membrane protein

Query Match . 11.1%; Score 295; DB 2; Length 391;
Best Local Similarity 26.9%; Pred. No. 7e-16;
Matches 98; Conservative 64; Mismatches 130; Indels 72; Gaps 16;

QY 75 YVEMTVGSPPTLNLVDTGSSNFAVGA----APHFTHRYORQLSTYRDLKRGVYV 130
Db 74 YFGTISIGSPPTNFVIFDTGSSNLWVPSVYCTSPACQTHPVHPSLSSTREVGNFSI 133
QY 131 PYTOGKWEGLGDLVSIPIHGPNTVTRANIAAITESDKFFINGSNNEGILGLAYAEIARP 190
Db 134 QYGTGSLTGLIGADQYSV-EGLTVVGOQFGESVQEPKTFVH-AEFDGILGLGSPSLAA- 190
QY 191 DDLSEPFDSLVKQTHVPNLFSLHLCGAGFPLNQSEVLASVGGSMIIGGDHSLYTGSLW 250
Db 191 -GGVTPVFDNMAQ-----NLVALPM-----FSVTMSSNPGSGSELTFGGYDFSHFGSLN 241
QY 251 YPIRREWEYEVIIVRVEINGQDLKMDCKEYNDKSIDVSGTNNLRLPKKVFEEAAVKS 310
Db 242 WPVTKQAYWQIALDGIQVG--DSVMFCSE--CCQAIVDTGTSLTGP---PGKIKQLQ 293
QY 311 AASSTKFPDGFVLGEOLVCWQAGTTPWNIFPVISLYLMGEVNTQSFRI-----TILPQ 365
Db 294 EALGATVDEGY----SVQC-----ANLNMMLDVT---FINGVPYTLNPTA 333
QY 366 YLRPVEDVATSDQDCYKFAISQSSTG-----TVMGAVIMEGFYVVDRAK 412
Db 334 Y--TLLDFVDMQVC-----STGEGLGIOPAGPLWILGDFVIRQFYAVFDRGN 383
QY 413 IGFA 416
Db 384 VGLA 387

Search completed: June 5, 2003, 16:17:40
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:09:05 ; Search time 84 Seconds
(without alignments)
1228.924 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLMGAGVLPAAH.....CLRLRQHQHDFADDISLLK 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2506.5	94.1	476	4 Q9BYC1	Q9byc1 homo sapien
2	2478.5	93.0	532	4 Q9ULS1	Q9uls1 homo sapien
3	2398	90.0	457	4 Q9BYC0	Q9byc0 homo sapien
4	2258.5	84.8	432	4 Q9BYB9	Q9byb9 homo sapien
5	1407	52.8	266	11 Q9CUU5	Q9cuu5 mus musculus
6	1155.5	43.4	439	4 Q9H2V8	Q9h2v8 homo sapien
7	1150	43.2	514	11 Q9JL18	Q9jll8 mus musculus
8	974.5	36.6	468	4 Q9NZL2	Q9nzl2 homo sapien
9	969.5	36.4	396	4 Q9NZL1	Q9nzl1 homo sapien
10	712.5	26.7	213	4 Q9P0D2	Q9p0d2 homo sapien
11	596.5	22.4	255	11 Q9R1P7	Q9rlp7 mus musculus
12	354.5	13.3	244	5 Q8WQY9	Q8wqy9 afrocaallis
13	335.5	12.6	391	5 Q9VKP6	Q9vxp6 drosophila
14	335	12.6	354	5 Q9GIX7	Q9gyx7 boophilus m
15	319	12.0	384	13 Q9DEC2	Q9dec2 xenopus lae
16	313.5	11.8	385	13 Q9DEC4	Q9dec4 rana catesb

17	312.5	11.7	386	6 Q9BGU5	Q9bgu5 bos taurus
18	311	11.7	387	6 Q9GMY8	Q9gmy8 sorex unqui
19	310	11.6	372	5 Q9VLK3	Q9vlk3 drosophila
20	308	11.6	386	6 Q9CMV7	Q9cmv7 rhinolophus
21	307.5	11.5	383	13 Q9DEC3	Q9dec3 xenopus lae
22	307.5	11.5	387	13 Q9DDV5	Q9ddv5 salvelinus
23	307	11.5	387	6 Q9GMY9	Q9gmy9 suncus muri
24	306.5	11.5	383	13 Q9DE45	Q9de45 salvelinus
25	305.5	11.5	376	13 Q9PUR8	Q9pur8 pseudopleur
26	305	11.4	384	13 Q91322	Q91322 rana catesb
27	304	11.4	382	13 Q9PRG9	Q9prg9 gallus gall
28	304	11.4	423	5 Q9VKP7	Q9vxp7 drosophila
29	298.5	11.2	386	6 Q9GMY6	Q9gmy6 canis famil
30	296.5	11.1	396	13 Q93428	Q93428 chionodraco
31	293.5	11.0	381	6 Q9GK11	Q9gk11 camelus dro
32	293	11.0	399	13 Q93458	Q93458 podarcis si
33	290.5	10.9	380	6 Q28950	Q28950 sus scrofa
34	289.5	10.9	399	13 Q9DD89	Q9dd89 brachydanio
35	287.5	10.8	444	5 Q21966	Q21966 caenorhabdi
36	287	10.8	427	5 P91802	P91802 schistosoma
37	286.5	10.8	378	13 Q9PUR9	Q9pur9 pseudopleur
38	284	10.7	389	13 Q9PWK1	Q9pwk1 gallus gall
39	284	10.7	398	13 P87370	P87370 oncorhynch
40	283	10.6	390	6 Q8SQ41	Q8sq41 canis famil
41	283	10.6	422	5 Q9G906	Q9g906 onchocerca
42	282.5	10.6	389	13 Q9W643	Q9w643 gallus gall
43	282.5	10.6	390	6 Q9GK10	Q9gk10 camelus dro
44	281.5	10.6	370	6 Q9TTW1	Q9ttw1 bos taurus
45	280.5	10.5	396	13 Q9DEX3	Q9dex3 clupea hare

ALIGNMENTS

RESULT 1

Q9BYC1 PRELIMINARY; PRT; 476 AA.
AC Q9BYC1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Beta-site APP cleaving enzyme I-476.
OS BACE.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H, Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
cleaving enzyme (BACE) and their effect on amyloid beta-peptide
production.";
RL Neurosci. Lett. 307:9-12(2001).
DR EMBL: AB050436; BAB40931.1; -
DR HSP: P32329; IYPS.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp; 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN1.
SQ SEQUENCE 476 AA; 52907 MW; 6C8C87F8A953AF66 CRC64;

Query Match 94.1%; Score 2506.5; DB 4; Length 476;
Best Local Similarity 95.0%; Pred. No. 6.9e-193;
Matches 476; Conservative 0; Mismatches 0; Indels 25; Gaps 1;

QY 1 MAQALPWLMLMGAGVLPAAHQTGHGIRLPLRSLGAGAPLGLRLPRETDEPEPEGRSGF 60
|||||
DB 1 MAQALPWLMLMGAGVLPAAHQTGHGIRLPLRSLGAGAPLGLRLPRETDEPEPEGRSGF 60
|||||

```

61 VEMVDNLRCKSGQYVVEWTVGSPQTNLNILVDGTGSNFVGAAPHFLHRYTORQLSST 120
|||||
61 VEMVDNLRCKSGQYVVEWTVGSPQTNLNILVDGTGSNFVGAAPHFLHRYTORQLSST 120
|||||
121 YRDLRGVVVPYTQGWEGELGTDLYSI PHGPNVTVRANIAAITESDKFTINGSNNEGIL 180
|||||
121 YRDLRGVVVPYTQGWEGELGTDLYSI PHGPNVTVRANIAAITESDKFTINGSNNEGIL 180
|||||
181 GLAYAEIARPDOSLEPFEDSLVKQTHVPLNFLSLHL CGAGFP LNQSEVLASVGGSMTIGGI 240
|||||
181 GLAYAEIAR-----LCGAGFP LNQSEVLASVGGSMTIGGI 215
|||||
241 DHSLYTCSLWYTPIRREWYEVII VRVEINGODLRKMDCKEYNIDKSIVDSGTTNLRLPKK 300
|||||
216 DHSLYTCSLWYTPIRREWYEVII VRVEINGODLRKMDCKEYNIDKSIVDSGTTNLRLPKK 275
|||||
301 VFEEAAVRSIKAAASTETKEFPDGFWLGBQLVCWAAGTTPWNIPFVISLYLMGEVWTNQSFRT 360
|||||
276 VFEEAAVRSIKAAASTETKEFPDGFWLGBQLVCWAAGTTPWNIPFVISLYLMGEVWTNQSFRT 335
|||||
361 ILPOOYLRPVEDVATSDDDCYKFALISQSSTGTVMGAV IMEGFYVDFRRKRIGFAVSAC 420
|||||
336 ILPOOYLRPVEDVATSDDDCYKFALISQSSTGTVMGAV IMEGFYVDFRRKRIGFAVSAC 395
|||||
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDESTLMTIAYVNAAICAFMLPCLCLMVQCW 480
|||||
396 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDESTLMTIAYVNAAICAFMLPCLCLMVQCW 455
|||||
481 RCLRLCROHQHDDFADDISLLK 501
|||||
456 RCLRLCROHQHDDFADDISLLK 476

```

[illegible]

Query Match 90.0%; Score 2398; DB 4; Length 457;
Best Local Similarity 91.0%; Pred. No. 3.4e-184;
Matches 456; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAQALPWLWAGAGVLPAGHTGIRLPLRSLGGLGAPLGLRLPRETDEPEPGRGSGF 60
DB 1 MAQALPWLWAGAGVLPAGHTGIRLPLRSLGGLGAPLGLRLPRETDEPEPGRGSGF 60
QY 61 VEMVDNLGRSGGGYVEMTVGSPPTLNILVDTGSSNFAGVGAAPHPFLHRYTORLSST 120
DB 61 VEMVDNLGRSGGGYVEMTVGSPPTLNILVDTGSSNFAGVGAAPHPFLHRYTORLSST 120
QY 121 YRDLRGVYVPTTGKWEGLGDLVSIHPGPNVTVRANTAAITESDKFFINGSNWEGIL 180
DB 121 YRDLRGVYVPTTGKWEGLGDLVSIHPGPNVTVRANTAAITESDKFFINGSNWEGIL 180
QY 181 GLAYAEIARDDSLPEFFDSLVKQTHVFNLSLHLCGAGPLNQSEVLASVGGSGMIIGGI 240
DB 146 -----PDDSLPEFFDSLVKQTHVFNLSLHLCGAGPLNQSEVLASVGGSGMIIGGI 196
QY 241 DHSLYTGLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNRLPKK 300
DB 197 DHSLYTGLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNRLPKK 256
QY 301 VFEAAVKSIIKAASSTKFPDGFGLGEOLVQWAGTTPWNIFFPVISLYLMGEVTVNQSPRIT 360
DB 257 VFEAAVKSIIKAASSTKFPDGFGLGEOLVQWAGTTPWNIFFPVISLYLMGEVTVNQSPRIT 316
QY 361 ILPQOYLRPVEDVATSDODCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSAC 420
DB 317 ILPQOYLRPVEDVATSDODCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSAC 376
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESLMTIAYVMAAICALFMLPLCLMVCOW 480
DB 377 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESLMTIAYVMAAICALFMLPLCLMVCOW 436
QY 481 RCLRLRQHQHDFADDSISLLK 501
DB 437 RCLRLRQHQHDFADDSISLLK 457

RESULT 4
Q9BYB9 ID Q9BYB9 PRELIMINARY; PRT; 432 AA.
AC Q9BYB9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-site APP cleaving enzyme I-432.
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production";
RL Neurosci. Lett. 307:9-12(2001).
DR EMBL; AB050438; BAB40933.1; -;
DR HSP; P32329; IYPS.
DR InterPro; IPR001461; AsproteaseA1.
DR InterPro; IPR001969; AsproteaseA1.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PRPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 432 AA; 4821 MW; 96FC81E6F0ED01B CRC64;

Query Match 84.8%; Score 2258.5; DB 4; Length 432;

Best Local Similarity 86.2%; Pred. No. 5e-173;
Matches 432; Conservative 0; Mismatches 0; Indels 69; Gaps 1;

QY 1 MAQALPWLWAGAGVLPAGHTGIRLPLRSLGGLGAPLGLRLPRETDEPEPGRGSGF 60
DB 1 MAQALPWLWAGAGVLPAGHTGIRLPLRSLGGLGAPLGLRLPRETDEPEPGRGSGF 60
QY 61 VEMVDNLGRSGGGYVEMTVGSPPTLNILVDTGSSNFAGVGAAPHPFLHRYTORLSST 120
DB 61 VEMVDNLGRSGGGYVEMTVGSPPTLNILVDTGSSNFAGVGAAPHPFLHRYTORLSST 120
QY 121 YRDLRGVYVPTTGKWEGLGDLVSIHPGPNVTVRANTAAITESDKFFINGSNWEGIL 180
DB 121 YRDLRGVYVPTTGKWEGLGDLVSIHPGPNVTVRANTAAITESDKFFINGSNWEGIL 145
QY 181 GLAYAEIARDDSLPEFFDSLVKQTHVFNLSLHLCGAGPLNQSEVLASVGGSGMIIGGI 240
DB 146 -----LCGAGPLNQSEVLASVGGSGMIIGGI 171
QY 241 DHSLYTGLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNRLPKK 300
DB 172 DHSLYTGLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNRLPKK 231
QY 301 VFEAAVKSIIKAASSTKFPDGFGLGEOLVQWAGTTPWNIFFPVISLYLMGEVTVNQSPRIT 360
DB 232 VFEAAVKSIIKAASSTKFPDGFGLGEOLVQWAGTTPWNIFFPVISLYLMGEVTVNQSPRIT 291
QY 361 ILPQOYLRPVEDVATSDODCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSAC 420
DB 292 ILPQOYLRPVEDVATSDODCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSAC 351
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESLMTIAYVMAAICALFMLPLCLMVCOW 480
DB 352 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESLMTIAYVMAAICALFMLPLCLMVCOW 411
QY 481 RCLRLRQHQHDFADDSISLLK 501
DB 412 RCLRLRQHQHDFADDSISLLK 432

RESULT 5
Q9CUU5 ID Q9CUU5 PRELIMINARY; PRT; 266 AA.
AC Q9CUU5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-site APP cleaving enzyme (Fragment).
GN BACE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=BRAIN;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK014390; BAB29317.1; -.
DR MEROPS; A01.004; -.
DR MGD; MGI:1346542; Bace.
DR InterPro: IPR001461; AspproteaseA1.
DR Pfam; PF00026; asp; 1.
FT NON_TER 1
SQ SEQUENCE 266 AA; 30201 MW; B913FDA8ADAB4238 CRC64;

Query Match 52.8%; Score 1407; DB 11; Length 266;
Best Local Similarity 98.9%; Pred. No. 5.9e-105;
Matches 263; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 236 IIGDHSLSYTGSLWYTPIRREWYEVIIYRVEINGODLKMCKEYNDKSIYVDSSTTNL 295
DB 1 IIGDHSLSYTGSLWYTPIRREWYEVIIYRVEINGODLKMCKEYNDKSIYVDSSTTNL 60

296 RLPKVFEEAAVKSIAKASSTKFKPDGFWLGEOLVCMQAGTTPWNIPFVISLYLMGEVTNQ 355
61 RLPKVFEEAAVKSIAKASSTKFKPDGFWLGEOLVCMQAGTTPWNIPFVISLYLMGEVTNQ 120

356 SFRITILPQOYLRPVEDVATSDQCYKFAISQSSTGTVMGAVIMEGFYVDFRARRKRGF 415
121 SFRITILPQOYLRPVEDVATSDQCYKFAISQSSTGTVMGAVIMEGFYVDFRARRKRGF 180

416 AVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPTDESTLMTIAYVMAAICALFMPPLCL 475
181 AVSACHVHDEFTAAVEGPFVTLDMEDCGYNIPTDESTLMTIAYVMAAICALFMPPLCL 240

476 MVCWRCLRCLRQHQHDFADDISLLK 501
241 MVCWRCLRCLRQHQHDFADDISLLK 266

RESULT 6
Q9H2V8 PRELIMINARY; PRT; 439 AA.
AC Q9H2V8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CDA13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=PHOCHROMOCYTOMA;
RA Li Y., Huang Q., Peng, Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF212252; AAG41783.1; -.
DR HSSP; P00797; 2REN.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 439 AA; 48275 MW; 02EC0E0E50F11602 CRC64;

Query Match 43.4%; Score 1155.5; DB 4; Length 439;
Best Local Similarity 49.9%; Pred. No. 2e-84;
Matches 219; Conservative 78; Mismatches 135; Indels 7; Gaps 4;

QY 63 MYDNLRGSGGYVEMTVGSPQPTLNLVDTGSSNFAVGAAPHPFLHRYQRQLSSTYR 122
DB 1 MYDNLRGSGGYVEMTVGSPQPTLNLVDTGSSNFAVGAAPHPFLHRYQRQLSSTYR 60

123 DLKRGVYVPYTGCKWEGELGDLVSLPHGPNVTVRANTAAITESDKFFINGSNWEGILGL 192

DB 61 SKGFDVTVKYGSTGTFVGEDLVTPKGFNTSFLVNIATIFESNFPLPGIKWNGILGL 120
QY 183 AYAEIARPDSDSLEPPFDLSLVKQTHVNLFLSLHLCGAGFPPLNOSEVLASVGGSMIIGDIDH 242
DB 121 AYATLAKPSSSLETFDLSLVTOANIPNVFSMQCGAGLPVAGS---GTNGGSLVGLGIEP 177
QY 243 SLATGSLWYTPIRREWYEVIIYRVEINGODLKMCKEYNDKSIYVDSSTTNLRLPKVYF 302
DB 178 SLYKGDWYTPIKREWYQIETLKLKLEIGGOSLNDCREYNADKAIYVDSGTTLLRLPKVYF 237
QY 303 EAAVKSIKASSTKFKPDGFWLGEOLVCMQAGTTPWNIPFVISLYLMGEVTNOSFRITIL 362
DB 238 DAVEAVARASLIPEFSDGEFTGSQLACWTNSSETPWSYFPKISIIYLRDENSRSFRITIL 297
QY 363 PQOYLRPVEDVATSDQCYKFAISQSSTGTVMGAVIMEGFYVDFRARRKRGFAVSACHV 422
DB 298 POLYIQPMWAGLNY-ECYRFGISFISNALVIGATVMEGFYVDFRAQKRVGFAASPCAE 356
QY 423 HDEFTAAVEGPFVTLDMEDCGYNIPTDESTLMTIAYVMAAICALFMPPLCLMVCOWR 481
DB 357 TAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVILLVLLPFR 416
QY 482 CLRCLRQHQHDFADDISLL 500
DB 417 CQR--RPRDPEVNVNDESSL 433

RESULT 7
Q9JLL18 PRELIMINARY; PRT; 514 AA.
AC Q9JLL18;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Aspartyl protease 1.
GN BACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
Choi D.K., Sugano S., Sakaki Y.;
RT "Molecular characterization of the mouse Aspl gene, a homolog of the
human Aspl (Down Syndrome Region aspartyl protease).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF216310; AAF36599.1; -.
DR HSSP; P32329; IYPS.
DR MEROPS; A01.041; -.
DR MGD; MGI:1860440; Bace2.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 514 AA; 55799 MW; A70725F2C1DF5B47 CRC64;

Query Match 43.2%; Score 1150; DB 11; Length 514;
Best Local Similarity 48.3%; Pred. No. 7e-84;
Matches 224; Conservative 76; Mismatches 144; Indels 20; Gaps 5;

QY 14 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRUPRETDEEPERGRGSEVMVDNLKSGQ 73
DB 43 ASAVPLGLTP---ELPRADGLA---LALPVRAT-----ANFLAVMDNLQDGSGR 86
QY 74 GYVEMTVGSPQPTLNLVDTGSSNFAVGAAPHPFLHRYQRQLSSTYRDLKRGVYVPY 133
DB 87 GYVLEMLTCTPQKQVILVDTGSSNFAVGAAPHPFLHRYQRQLSSTYRDLKRGVYVPY 146
QY 134 QCKWEGELGDLVSLPHGPNVTVRANTAAITESDKFFINGSNWEGILGLAYEATRPDDS 193

Db 147 QGWTGFGVDELVTIPKGFNSFLVNITATIFESNFFLPGLKNGIILGLAYAAALAKPSS 206
Qy 194 LEPPFDSLVLKQTHVPLNLSLHLCAGPPLNOSVELASVGGSMIIGDHSLSYTGSLWTP 253
Db 207 LETFDSILVAQAQIPDIFSMQCGAGLPVAGS---GTNGGSLVLGGIEPSLYKGDWYTP 263
Qy 254 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSTVDSGTTNLRPLPKVFEAAVKSIAAS 313
Db 264 IKEWYQIEILKLEIGQNLNLDREYNADKAVDSGTTLLRLPKQVDFVAVAVARTS 323
Qy 314 STEKFPDGLGEOLVCWQAGTTPWNIFFPVLSLYLMGEVNTNQSFRITLPOQLRPVEDV 373
Db 324 LIPEFGWFGAQLACTWSETPWAYFPKISIVLRDENASRSFRITLPOLYIQPMGA 383
Qy 374 ATSDQDCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSACHVHDEFRTAAVEG 433
Db 384 GFNY-ECYRGISSTNALVIGAVMEGFYVDFRARRVGFVSPCAEIEGTVSEISG 442
434 PFVTLMDCCGYNIPTQDESTLMTIAYMAAICAFMLPLCLMV 477
443 PFSTEDIASNCVPAQALNEPTLWISYALMSVCGAILLVLL 486

RESULT 8

Q9NZL2
ID Q9NZL2 PRELIMINARY; PRT; 468 AA.
AC Q9NZL2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aspartyl protease.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase."
RL Cyogenet. Cell Genet. 89:177-184(2000).
DR EMBL; AF188276; AAF35835.1; -
DR HSSP; P00797; 2REN.
DR MEROPS; A01.041; -
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPFIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 468 AA; 50324 MW; 717E0920126A0142 CRC64;

Query Match 36.6%; Score 974.5; DB 4; Length 468;
Best Local Similarity 40.5%; Pred. No. 7.6e-70;
Matches 210; Conservative 76; Mismatches 150; Indels 83; Gaps 10;

Qy 2 AQAALPWLKLM---GAGVLPAGHTQHGLRLPLRSGLG-----GAPL-----GLR 42
Db 7 ALLPLLAQMLRAAPLAPPT-----LPLRVAATNRVAVPTPGGTPAERHADGLA 61
Qy 43 LPRETDEPEEPGRGSGFVEMVDNLGRKSGGYVEMTVPQTNLILVDTGSSNFAVG 102
Db 62 LALE--PALASPAAGANFLAMVDNLQDSCGRGYLEMLIGTPPKQLIILVDTGSSNFAV 119
Qy 103 AAPHPFLHRYQROLSTYRDLRGVYVPTQGWEGELGTDLVSIHPGNVTVRANAA 162
Db 120 GTPHSYIDTVFDTSTSTYRSKGFDTVKYVQTGSGWTFVGDVLTIPKGFNTSLVNIAT 179
Qy 163 ITESDKFEFNGSNEGILGLAYAEIARPDOSLEPFFDSLVLKQTHVPLNLSLHLCAGPPL 222
Db 180 IFESNFFLPGLKNGIILGLAYATLAKPSSSLETFDFSLVTVQANIPNVFSNQMGAGLPV 239

Qy 223 NQSEVLASVGGSMIIGDHSLSYTGSLWYTPIRREWYEVIIIVRVEINGQDLKMDCKEYN 282
Db 240 AGS---GTNGGSLVLGGIEPSLYKGDWYTPKEWYQIEILKLEIGQSLNLDREYN 296
Qy 283 YKSIDVSGTTNLRPLPKVFEAAVKSIAASSTKEKFPDGLGEOLVCWQAGTTPWNIFF 342
Db 297 ADKAVDSGTTLLRLPKQVDFVAVAVARASLL----- 329
Qy 343 VISLYLMGEVNTNQSFRITLPOQLRPVEDVATSDQDCYKFAISOSSTGTVMGAVIMEGF 402
Db 330 -----YIQPMMGAGLNY-ECYRGISSTNALVIGAVIMEGF 365
Qy 403 YVDFRARKRIGFAVSACHVHDEFRTAAVEGFVTLMDCCGYNIPTQDESTLMTIAYVM 462
Db 366 YVDFRARKRIGFAVSACHVHDEFRTAAVEGFVTLMDCCGYNIPTQDESTLMTIAYVM 462
Qy 463 AAIC-ALFMLPLCLMVCMOWRCLRQOHHDDFADDSLL 500
Db 426 MSVCGAILLVLLVLLLPFCQR--RPRDPEVWDESSL 462

RESULT 9

Q9NZL1
ID Q9NZL1 PRELIMINARY; PRT; 396 AA.
AC Q9NZL1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aspartyl protease.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase."
RL Cyogenet. Cell Genet. 89:177-184(2000).
DR EMBL; AF188277; AAF35836.1; -
DR HSSP; P00797; 2REN.
DR MEROPS; A01.041; -
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPFIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 396 AA; 43013 MW; 5023A7AF391CEAC9 CRC64;

Query Match 36.4%; Score 969.5; DB 4; Length 396;
Best Local Similarity 49.3%; Pred. No. 1.5e-69;
Matches 200; Conservative 56; Mismatches 111; Indels 39; Gaps 9;

Qy 2 AQAALPWLKLM---GAGVLPAGHTQHGLRLPLRSGLG-----GAPL-----GLR 42
Db 7 ALLPLLAQMLRAAPLAPPT-----LPLRVAATNRVAVPTPGGTPAERHADGLA 61
Qy 43 LPRETDEPEEPGRGSGFVEMVDNLGRKSGGYVEMTVPQTNLILVDTGSSNFAVG 102
Db 62 LALE--PALASPAAGANFLAMVDNLQDSCGRGYLEMLIGTPPKQLIILVDTGSSNFAV 119
Qy 103 AAPHPFLHRYQROLSTYRDLRGVYVPTQGWEGELGTDLVSIHPGNVTVRANAA 162
Db 120 GTPHSYIDTVFDTSTSTYRSKGFDTVKYVQTGSGWTFVGDVLTIPKGFNTSLVNIAT 179
Qy 163 ITESDKFEFNGSNEGILGLAYAEIARPDOSLEPFFDSLVLKQTHVPLNLSLHLCAGPPL 222
Db 180 IFESNFFLPGLKNGIILGLAYATLAKPSSSLETFDFSLVTVQANIPNVFSNQMGAGLPV 239

223 NQSEVLASVCGSMIIGIDHSYLTSLWTPTRREWEYEVIIIVRVEINGDQKMDCKEYN 282
 240 AGS---GTNGSGLVIGIEPSLTKGDIWTPPKKEWYQIEILKLEIGQSLNDCREYN 296
 283 YKSDVSGTNTLRPKKPEAAVSIKAASSTKFPDGFGLGQVLCVQAGTTWNIFP 342
 297 ADKAIVDSGTTLLRLPKVDFEAVAVARASLIPEESDGFGLGQVLCVQAGTTWNIFP 356
 343 VISLYLMGEVNTNSFRITILPOQYLPRVEDVATSDQDCYKF-AISO 387
 357 KISYLRDENSRSFRITILPOK-LRVLQ-----CLKFPLGSLQ 393

RESULT 10
 ID Q9PD02 PRELIMINARY; PRT; 213 AA.
 AC Q9PD02;
 DT 01-OCT-2000 (TremBLrel. 15, Created)
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DT HSPCL04 (Fragment).
 Homo sapiens (Human).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CORD BLOOD;
 RA Zhang Q.H., Ye M., Zhou Y., Wu X.Y., Guan Z.Q., Wang L.,
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
 RT "Human partial CDS cloned from cd34+ stem cells."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF161367; AAF28927.1;
 DR InterPro: IPR001461; AspproteaseA1.
 DR Pfam: PF00026; asp; 1.
 FT NON_TER 1
 SQ SEQUENCE 213 AA; 24338 MW; EC9D3FA31CFA835C CRC64;
 Query Match 26.7%; Score 712.5; DB 4; Length 213;
 Best Local Similarity 83.5%; Pred. No. 2.6e-49;
 Matches 137; Conservative 4; Mismatches 12; Indels 11; Gaps 1;
 QY 238 GGDHSYLTSLWTPTRREWEYEVIIIVRVEINGDQKMDCKEYNVDKSGTNTLRL 297
 DB 1 GGDHSYLTSLWTPTRREWEYEVIIIVRVEINGDQKMDCKEYNVDKSGTNTLRL 60
 QY 298 PKVFEAAVSIKAASSTKFPDGFGLGQVLCVQAGTTWNIFPVISLYLMGEVNTNSF 357
 61 PKVFEAAVSIKAASSTKFPDGFGLGQVLCVQAGTTWNIFPVISLYLMGEVNTNSF 120
 358 RITILPOQLRPVEDVATSDQDCYKF-AISOSSSTGTVMGAVIMBG 401
 121 RITILPOQLRP-----WKMPRPKTTVTCHLTVIHG 153

RESULT 11
 ID Q9R1P7 PRELIMINARY; PRT; 255 AA.
 AC Q9R1P7;
 DT 01-MAY-2000 (TremBLrel. 13, Created)
 DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE Aspartyl protease (Fragment).
 GN BACE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A.
 RA Accorino M., Fumagalli P., Taramelli R., Ottolenghi S.;
 RT "Cloning of a gene from chromosome 21 Down region encoding a potential transmembrane protease."
 RT

Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF051150; AAD45964.1;
 DR MEROPS; A01.041;
 DR MGD; MGI:1860440; Bace2.
 DR InterPro: IPR001461; AspproteaseA1.
 DR InterPro: IPR001969; Aspprotease_site.
 DR PRINTS; PR00792; PPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
 KW Protease.
 FT NON_TER 1
 SQ SEQUENCE 255 AA; 28685 MW; 53DE317815996D63 CRC64;
 Query Match 22.4%; Score 596.5; DB 11; Length 255;
 Best Local Similarity 47.8%; Pred. No. 7e-40;
 Matches 109; Conservative 44; Mismatches 74; Indels 1; Gaps 1;
 QY 250 WYTPTRREWEYEVIIIVRVEINGDQKMDCKEYNVDKSGTNTLRLPKKVFEEAAVSI 309
 DB 1 WYTPKEEYQIEILKLEIGQNLNDCREYNADKALVDSGTTLLRLPKVDFAVVEAV 60
 QY 310 KAASSTKFPDGFGLGQVLCVQAGTTWNIFPVISLYLMGEVNTNSFRITILPOQLRP 369
 DB 61 ARTSLIPEESDGFGLGQVLCVQAGTTWNIFPVISLYLMGEVNTNSFRITILPOQLRP 120
 QY 370 VEDVATSDQDCYKFALISOSSSTGTVMGAVIMBGFYVDFRKRIGFAYVACHVHDEFRTA 429
 DB 121 MMGAGFNY-ECYRFGISSSTNALVIGATVMEGFYVDFRKRIGFAYVACHVHDEFRTA 179
 QY 430 AVEGPPVTLMEDCGYNIPQTDSTLMTIAYVMAICAFMLPLCLMV 477
 DB 180 EISGPPSTEDIASNCVPAQALNEPIWIVSYALMSVCGAILLVILL 227
 RESULT 12
 ID Q8WQY9 PRELIMINARY; PRT; 244 AA.
 AC Q8WQY9;
 DT 01-MAR-2002 (TremBLrel. 20, Created)
 DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE Aspartate protease (Fragment).
 GN APP.
 OS Aphrocallistes vastus.
 OC Eukaryota; Metazoa; Porifera; Hexactinellida; Hexasterophora;
 OC Lysacnoida; Rosellidae; Aphrocallistes.
 OX NCBI_TaxID=83887;
 [1]
 RN SEQUENCE FROM N.A.
 RA Mueller W.E.G., Mueller I.M., Grebenjuk V.A.;
 RT "Urmatazoa: Origin and evolution of the common ancestor of Metazoa."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ304863; CAC83293.1;
 DR InterPro: IPR001461; AspproteaseA1.
 DR InterPro: IPR001969; Aspprotease_site.
 DR Pfam: PF00026; asp; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
 DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
 KW Protease.
 FT NON_TER 244
 SQ SEQUENCE 244 AA; 26366 MW; 6536902661E0E4C7 CRC64;
 Query Match 13.3%; Score 354.5; DB 5; Length 244;
 Best Local Similarity 35.6%; Pred. No. 1.8e-20;
 Matches 80; Conservative 41; Mismatches 91; Indels 13; Gaps 6;
 QY 64 VDNLRCKSGQYVEMTVGSPDPTNLIVDTGSSNFSSFAVGAAPHPLHRY--YQQLSSTY 121
 DB 22 VYQLOQPSGYSYLVNIGTTPQEFKVLVDSSSNFAVAAKFGYQDYPRFDSLSKTF 81
 QY 122 RDLRGVYVPTQGWKEGELGDLVS----TPHGPNVIVRANIAI--TESDKFFLINGSW 176
 DB 82 RDINSEGVKVIDGWSGVRGDEYFAFASDVTTNASKSVKVVYSLIESVSEGFETSGW 141

Qy 177 EGILGLAYAEIARDDSLPEFDSVLKQTHVNPFLSHLCGAGFPFNQSE--VLASVGSGM 233
| | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 142 VGILGMGYAVLAKPSSITTPVMDLSVSGVTSDKFGLQLCQ---PLNSDRLELVNOGGKM 198
| | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 236 IIGIDHSLYTGCS--LWYTPIRREMYEYVIIVRVEINCOGLKMCDC 278
| | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 199 SLGSARBLPSNDSOVFRVAITEBSFYEVILTNIKVGSTSIDLPCC 243
| | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 13

ID Q9VKP6 PRELIMINARY; PRT: 391 AA.

AC Q9VKP6;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE CG17134 protein (RE02351p).

GN CG17134

[1]

Drosophila melanogaster (Fruit fly).

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Phrygoryga; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.W., Hoskins R.A., Galle R.F.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Randall R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berham B.P., Bhargava D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwu C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Swirskas R., Tector C., Turner R., Venter E.H., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of *Drosophila melanogaster*.";
Science 287:2185-2195(2000).
[2]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
Miranda A., Mundall C.J., Nunco J., Pacle J., Paragas V., Park S.,

RA	Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA	Celniker S.;
RA	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC	EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR	EMBL; AE003630; AAF53016.1; -
DR	EMBL; AY070911; AAL48533.1; -
DR	HSSP; P00794; 4CMS.
DR	FlyBase; FBgn0032304; CG17134.
DR	InterPro; IPR001461; Asparticase1.
DR	InterPro; IPR001969; Aspprotease_site.
DR	Pfam; PF00026; asp.1.
DR	PRINTS; PR00792; PPSIN.
DR	PROSITE; PS00141; ASP_PROTEASE; 2.
KW	Aspartyl protease; Hydrolase.
SQ	SEQUENCE 391 AA; 42016 MW; 99A18E7131025E1B CRC64;

Query Match	12.6%;	Score 335.5;	DB 5;	Length 391;
Best Local Similarity	29.1%;	Pred. NO. 1.2e-18;		
Matches 105;	Conservative 62;	Mismatches 145;	Indels 49;	Gaps 13

Qy	65	DNLGRKSGQGYVEMTVGSPQTLNILDVTSSNFAVGAAPHF-----LHRYQORQLS 118
Db	66	ENLHNSNNYYGVIAIGTPQQRNILDTSANLWVPSASCPASNTACQRHINKYDSSAS 125
Qy	119	STYRDLARKVVPVYQTQKWEGLGTLVSIHPGPNVTVRANIAATESDKFFINGSNWEG 178
Db	126	STVYANGEEPALEYCTGSLGFLSNDIVTIA-GISIQNTGCEALSEPCTTFVD-APFAG 183
Qy	179	ILGLAYAEIARPDLSLEPFFDSLVKQTHVPN-LFSLHLCGAGFPLNQSEVLASVGSMII 237
Db	184	ILGLAFSAIA--VDGVTVPFDNMISQGLDPEVISFLYLRQG-----TAVRGCELIL 233
Qy	238	GGIDHSLYTGSLSWTPIRREKYYEVIIVRVEINGODLKMDCKEYNYDKSIVDSGTTNLR 297
Db	234	GGIDSSLYRGSLLTVPVSVPAYWQFKVNTIKTNGTLLCNGC-----QAIADTGTSLIA 287
Qy	298	PKKVFEAAVRSIKAASSTKFPDGFGEQLV-CWQAGTTPWNIFPVISLYLMGEVNTQS 356
Db	288	PLAAYRKINROLGATDND-----GEAFVRCGRVSS-----LPKVNINIGGTV---- 329
Qy	357	FRITLFOQYLPRVEDVATSDODCYK-PAISOSSTGTWNGAVIMGEFVYVFDRAKRIGF 415
Db	330	--FPLAPRDYI--VKVTQNGQTYCMASFTYMEGLSFWILGDVFIGKFYTFDKNGERIGF 385
Qy	416	A 416
Db	386	A 386

RESULT 14	
Q9GYX7	PRELIMINARY; PRT; 354 AA.
ID	Q9GYX7
AC	Q9GYX7;
DT	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Heme-binding aspartic proteinase (Fragment).
OS	Boophilus microplus (Cattle tick).
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC	Parasitiformes; Ixodida; Ixodidae; Boophilidae; Boophilus.
OX	NCBI_TaxID=6941;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=PORTO ALEGRE; TISSUE=OVARY;
RA	Sorgine M.H.F., Logullo C., Zingali R.B., Paiva-Silva G.O.,
RA	Juliano L., Oliveira P.L.;
RT	"A heme-binding aspartic proteinase from the eggs of the hard tick
RT	Boophilus microplus.;"
RL	J. Biol. Chem. 0:0-0(2000).
DR	EMBL; AF286865; AAG00993.1; -
DR	HSSP; P00797; 2REN.
DR	MEROPS; A01.054; -

DR InterPro: IPR001461; AspproteaseA1.
 DR Pfam: PF00026; asp: 1.
 DR PRINTS: PR00792; PEPsin.
 FT NON_TER 1
 SQ SEQUENCE 354 AA; 38512 MW; B50D7C638CF27091 CRC64;
 Query Match 12.6%; Score 335; DB 5; Length 354;
 Best Local Similarity 26.2%; Pred. No. 1.2e-18;
 Matches 101; Conservative 75; Mismatches 134; Indels 76; Gaps 15;
 QY 64 VDLNRK-----SGGYVEMTVGSPPTLNLVDTGSSNFAVGAAPHPFL- 109
 DB 13 VTEIRGALGDPITPILITNVMQFYGIITCTPQSFLLMDTGSSNEW-----PSTN 67
 QY 110 -----HRYQROLSTYRDLRGVVPYTOGKWEGLGTDLSVIPHGNVTVRANI 160
 DB 68 CDOSMACROHAKYDSSKSTFTKSGRIIRIRVSGGVVRGITSIDNVG--GRATVTOYKF 125
 QY 161 AATSDKFFINGSNWEGILGLAYAEIARPDPSLEFFDSLVKQTHVPN-LFSLHLCGAG 219
 DB 126 AEMHSDGKLFRAKYDGIPLAFPSISQ--NNQLPLFDAMVKQGVVRQAVESLYL--SK 181
 QY 220 FPLNQSEVLASVGGMIIGDHSITGSLWYTPIRREWYEVIIIVRVEINGODLKM-DC 278
 DB 182 QPSEQN-----GGEYFGGINAQRYTGAHYVPSQAAHWOVMDNINNVQGTTLCCVGC 235
 QY 279 KEYNDKSIDVSGTTLNLRPKKVFEEAAVKSIIKAASSTEFKPDGFWLGEOLVCWQAGTTPW 338
 DB 236 -----PTVDSGTSFSLGP-----SADVELNRVIGATTAAGY-----FEVNCAT 277
 QY 339 NIFPVISLYLMGE---VTNQSFRITILPOQYLRPVEDVATSDQDCYKFAISQSTGT--- 392
 DB 278 SSLRPITENLNGKSFPLQGEAYTIRI-----PLTTGGQCFTRISESDASGTNLW 327
 QY 393 VMGAVIMEGYVVDRAKRRIGFAVS 418
 DB 328 ILGAVETOTYTYVFDRAQNRVGPATA 353

RESULT 15

Q9DEC2 Q9DEC2 PRELIMINARY; PRT; 384 AA.
 AC Q9DEC2;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Pepsinogen A.
 OS Xenopus laevis (African clawed frog).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ikuzawa M., Inokuchi T., Kobayashi K., Yasumasu S.;
 RT "Molecular Cloning of Pepsinogens in Adult Xenopus laevis and Bullfrog
 RT Rana catesbeiana";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 DR EMBL: AB045380; BAB20798.1;
 DR HSSP: P00790; IPSN.
 DR MEROPS: A01.001;
 DR InterPro: IPR001461; AspproteaseA1.
 DR InterPro: IPR001969; Aspprotease_site.
 DR Pfam: PF00026; asp: 1
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 1.
 KW Aspartyl protease; Hydrolase.
 SQ SEQUENCE 384 AA; 41888 MW; 729DD2E7953D9072 CRC64;
 Query Match 12.0%; Score 319; DB 13; Length 384;
 Best Local Similarity 27.0%; Pred. No. 2.5e-17;

Matches 118; Conservative 70; Mismatches 163; Indels 86; Gaps 21;
 QY 8 LLLWMGAGVLPAGHQHGRPLRLPSRLGLG-----APLGLRLPRETDEEPEECRR----- 57
 DB 3 LLLLLGLVVL-----SECYVVPRLKGESFRNRLQRLGL-LGDYLLKKNYPNASKYFPTLA 57
 QY 58 GSFVEMVNLRGKSGQGYVEMTVGSPPTLNLVDTGSSNFAVGAAPHPFL-----H 110
 DB 58 OSSAETLQNYMDIE---YVGITISCTPPQEFVTFDGTGSANLW---PSVYSSQACSNNH 111
 QY 111 RYQQLSSTYRDLRGVVPYTOGKWEGLGTDLSVIPHGNVTVRANIAITESDK-F 169
 DB 112 NRENFQSQSTFOATNTPVSIQYGTGSMGFLGYDTLQV---CNIQISNMFLSESEPGS 168
 QY 170 FINGSNWEGILGLAYAEIARPDPSLEFFDSLVKQTHVP-NLFLSHLCGAGFPLNQSEVL 228
 DB 169 FIYYSPFDGILGLAFPSIA--SSQATPVFDNWSOGLIPQNLFSVYLSDDG----- 217
 QY 229 ASVGGSMIIGDHSITGSLWYTPIRREWYEVIIIVRVEINGODL--KMDCKEYNYDKS 286
 DB 218 -QTGSVVLFGGVDSYSGSLNNVPLTAETWQITLDSVINGQVIACQSC-----QA 270
 QY 287 IYDSGTTNLRPKKVFEEAAVKSIIKAASSTEFKPDGFWLGEOLV-CWQAGTTPNIFPVIS 345
 DB 271 IYDTGSLMTGPSTPI-ANIQYIGASQDSN-----GQYVINCNNISNMPTIIVF----- 318
 QY 346 LYLMGEVNTNQSFRITILPOQYLRPVEDVATSDQDCYK-FAISQSTGT---VMGAVIM 399
 DB 319 -----TINGVQYPLSPSAYVRQNOGCGSGFOAMNLPNTNSGDLWILGDVFI 364
 QY 400 EGYVYVDFRARRKIGFA 416
 DB 365 RQYFTVDFRANNVVAIA 381

Search completed: June 5, 2003, 16:14:17
 Job time : 87 secs